

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
1 March 2001 (01.03.2001)

PCT

(10) International Publication Number  
**WO 01/14421 A1**

(51) International Patent Classification<sup>7</sup>: C07K 14/315,  
C12N 15/31, C07K 16/12, A61K 39/09, 39/40

(21) International Application Number: PCT/US00/23417

(22) International Filing Date: 25 August 2000 (25.08.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
60/150,750 25 August 1999 (25.08.1999) US

(71) Applicant: MEDIMMUNE, INC. [US/US]; 35 West  
Watkins Mill Road, Gaithersburg, MD 20878 (US).

(72) Inventors: KOENIG, Scott; 10901 Ralston Road,  
Rockville, MD 20852 (US). HEINRICHS, Jon; 9 Peach

Leaf Court, North Potomac, MD 20878 (US). JOHNSON,  
Leslie, Sydnor; 20147 Laurel Hill Way, Germantown, MD  
20874 (US). ADAMOU, John, E.; 20822 Shamrock Glen  
Circle, Germantown, MD 20874 (US).

(74) Agents: GRANT, Alan, J. et al.; Carella, Byrne, Bain,  
Gilfillan, Cecchi, Stewart & Olstein, 6 Becker Farm Road,  
Roseland, NJ 07068 (US).

(81) Designated States (*national*): AL, AM, AT, AU, AZ, BA,  
BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM,  
EE, ES, FI, GB, GD, GE, GH, GM, HU, ID, IL, IN, IS, JP,  
KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA,  
MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU,  
SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ,  
VN, YU, ZA, ZW.

(84) Designated States (*regional*): ARIPO patent (GH, GM,  
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian  
patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European

[Continued on next page]

(54) Title: HOMOLOGS OF A PNEUMOCOCCAL PROTEIN AND FRAGMENTS FOR VACCINES

*S. aureus* 6390  
*S. pyogenes* M1  
*S. pyogenes* M6  
*S. pneumoniae*

B P B P B P B P λ



(57) Abstract: The invention is directed to isolated polypeptides bearing sequence homology to the Sp36 protein found in pneumococcal organisms, such as *Streptococcus pneumoniae*. Polynucleotides encoding such polypeptides are also disclosed. The invention also relates to antibodies specific for the disclosed polypeptides and to uses of such antibodies in the treatment of diseases caused by staphylococci as well as group A and B streptococci. In addition, the invention relates to the use of the disclosed polypeptides in compositions and as vaccines and for prophylactic uses such as in vaccination of animals, especially humans, against a wide variety of streptococcal, staphylococcal and other diseases.

WO 01/14421 A1



patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE,  
IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG,  
CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

— Before the expiration of the time limit for amending the  
claims and to be republished in the event of receipt of  
amendments.

**Published:**

— With international search report.

For two-letter codes and other abbreviations, refer to the "Guid-  
ance Notes on Codes and Abbreviations" appearing at the begin-  
ning of each regular issue of the PCT Gazette.

## HOMOLOGS OF A PNEUMOCOCCAL PROTEIN AND FRAGMENTS FOR VACCINES

5

This application claims the priority of U.S. Provisional Application 60/150,750, filed August 25, 1999, the disclosure of which is hereby incorporated by reference in its entirety.

10

## FIELD OF THE INVENTION

This invention relates generally to the field of bacterial antigens and their use, for example, as immunogenic agents in humans and animals to stimulate an immune response. More specifically, it relates to the vaccination of mammalian species, especially humans, with one or more polypeptides derived from gram positive bacteria and which show sequence homology with an immunogenic polypeptide obtained from *Streptococcus pneumoniae*.

20

## BACKGROUND OF THE INVENTION

Polypeptides derived from gram positive bacteria are useful for stimulating production of antibodies that protect the vaccine recipient against infection by a wide range of serotypes of pathogenic gram positive bacteria, including *S. pneumoniae*. Further, the invention relates to antibodies against such polypeptides useful in diagnosis and passive immune therapy with respect to diagnosing and treating such pneumococcal infections.

30

The genus *Streptococcus* contains a variety of species responsible for causing disease in mammals, including humans, while also encompassing species that constitute normal flora in humans and other mammals. Among the bacterial species implicated in the etiology of diseases in humans are *S. pyogenes* (part of the group A streptococcal bacteria, herein designated "GAS" for "group A streptococci"), *S. pneumoniae* (referred to as "pneumococcus") and *S. agalactiae* (the group B streptococci or "GBS"). The group A streptococci cause serious diseases such as necrotizing fasciitis, scarlet fever and sepsis, as well as less virulent diseases such as impetigo and pharyngitis. The pneumococci are the most common cause of community-acquired pneumonia and are also responsible for more than half of all cases of otitis media in children. The pneumococci are also the second most common pathogen associated with bacterial meningitis. The group B streptococci are the most prevalent pathogen associated with illness and death among newborns in the United States.

Currently, there are no vaccines available for the prevention of diseases caused by the group A and group B streptococci and presently available pneumococcal vaccines are not effective in children under 2 years of age or in the elderly due to the poor immunogenicity of the capsular carbohydrates that compose the current vaccine. It would therefore be highly advantageous to produce a vaccine that would prevent infection by these classes of pathogen, especially in the age groups mentioned.

In addition to the pathogens just described, some bacteria of the genus *Staphylococcus* are also of clinical importance. In fact, two of these are among the leading causes of nosocomial infections (infections acquired while in the hospital). Both *Staphylococcus aureus* and *Staphylococcus epidermidis* readily colonize the skin of healthy individuals and can cause acute disease in patients following immunosuppression or traumatic injury. Infections caused by these species include bacteremia, endocarditis,

osteomyelitis, wound infections and infections associated with indwelling catheters.

*Streptococcus pneumoniae* is a gram positive bacterium that is a major causative agent in invasive infections in animals and humans, such as the aforementioned sepsis, meningitis, and otitis media, as well as lobar pneumonia (Tuomanen, *et al. New England J. of Medicine* 322:1280-1284 (1995)). As part of the infection process, pneumococci readily bind to non-inflamed human epithelial cells of the upper and lower respiratory tract by binding to eukaryotic carbohydrates in a lectin-like manner (Cundell *et al., Micro. Path.* 17:361-374 (1994)). Conversion to invasive pneumococcal infections for bound bacteria may involve the local generation of inflammatory factors which may activate the epithelial cells to change the number and type of receptors on their surface (Cundell, *et al., Nature*, 377:435-438 (1995)). Apparently, one such receptor, platelet activating factor (PAF) is engaged by the pneumococcal bacteria and within a very short period of time (minutes) from the appearance of PAF, pneumococci exhibit strongly enhanced adherence and invasion of tissue. Certain soluble receptor analogs have been shown to prevent the progression of pneumococcal infections (Idanpaan-Heikkila *et al., J. Inf. Dis.*, 176:704-712 (1997)). A number of other proteins have been suggested as being involved in the pathogenicity of *S. pneumoniae*.

*Streptococcus pneumoniae* itself has been shown to contain a gene which encodes a protein designated herein as Sp36. This protein has a predicted molecular mass of 91,538 Da and contains 5 histidine triad motifs (proposed to be involved in metal binding). The gene encoding this protein appears to be present the 23 serotypes comprising the current commercially available pneumococcal-capsular vaccine. Immunization of mice with this protein, in the presence of Freund's adjuvant, stimulates an immune

response which protects these mice from an intraperitoneal challenge with a dose of virulent pneumococci that would normally kill the mice.

For the reasons already stated above, there not only remains a need  
5 for identifying polypeptides having epitopes in common from various strains of *S. pneumoniae* but also from a broader spectrum of gram positive bacteria in order to utilize such polypeptides as vaccines to provide protection against a wide variety of infectious organisms.

## 10 BRIEF SUMMARY OF THE INVENTION

In accordance with the present invention, there is provided vaccines that include polypeptides obtained from gram positive bacteria other than *S. pneumoniae*, as well as variants of said polypeptides and active fragments  
15 of such polypeptides.

The present invention is also directed to novel genes, and the polypeptides encoded thereby, derived from gram positive bacteria other than *S. pneumoniae*, and which bear sequence homology to the Sp36 gene  
20 already described. Such gram positive bacteria include the group A and B streptococci, as described herein, as well as species of the genus *Staphylococcus*, especially *S. aureus*.

In a particular embodiment, the present invention is directed to  
25 specific gene sequences, and proteins encoded thereby, derived from the group A and group B streptococci, and to the use of such expressed polypeptides and proteins as the basis for pharmaceutical compositions useful as vaccines and as a means for enabling isolation of antibodies with therapeutic and/or prophylactic activity (such as would be useful in  
30 preparing products like CytoGam).

In a further embodiment, the present invention also relates to the preparation and use of fragments of the novel polypeptides disclosed her in, such fragments being immunogenic in nature and being useful in the preparation of vaccines against diseases caused by the pathogens from which such polypeptides, and fragments thereof, are derived.

10

### BRIEF DESCRIPTION OF DRAWINGS

Figures 1 shows the results of a Southern blot of genomic DNA from *S. aureus*, *S. pyogenes*, and pneumococcus. The DNA was digested with restriction nucleases *Bam*HI or *Pvu*II, and after electrophoresis and transfer to a nylon membrane, was probed with a labeled DNA fragment encompassing the pneumococcal gene encoding Sp36. The hybridization and washes were carried out under low stringency conditions. The results show hybridization by the labeled probe to a *S. aureus* fragment in both the *Bam*HI and *Pvu*II lanes and to two fragments in the *Pvu*II digests of two strains of *S. pyogenes*.

Figures 2 shows an alignment between the Sp36 amino acid sequence from *S. pneumoniae* strain N4 and the homologous sequences from *S. pyogenes* and *S. agalactiae*. Amino acids identical to those of the polypeptide from *S. pneumoniae* are boxed.

Figure 3 shows the results of a Southern blot of genomic DNA from *S. pyogenes*, *S. agalactiae*, and *S. pneumoniae* probed with DNA encoding the full length Sp36 homolog from *S. pyogenes*. The

hybridization was carried out under low stringency conditions. These results demonstrate that the *S. pyogenes* Sp36 homolog, used as a probe, is capable of detecting a homologous gene in *S. agalactiae* and pneumococcus.

5

Figure 4 shows the results of a western blot using rabbit polyclonal antiserum generated against recombinant Sp36 protein cloned from *S. pneumoniae* strain Norway 4. The results demonstrate that this antiserum not only reacts with the protein against which it was raised (here, Sp36),  
10 as well as to a protein of similar size in a lysate of a serotype 6B strain of pneumococcus, but also reacts with a recombinant protein encoded by the Sp36 homolog gene of group B streptococci.

Figure 5 shows the amino acid sequence for the GAS36 homologs  
15 with the histidine triad regions underlined (Fig. 5(a) and (b)) and the sequence for a GBS36 homolog (Fig. 5(c)) with its histidine triad regions underlined.

20

## DETAILED DESCRIPTION OF THE INVENTION

The present invention is directed to novel polynucleotides and polypeptides derived from species of gram positive bacteria, especially group A and B streptococci, and including the genus *Staphylococcus*, most  
25 especially *S. pyogenes* (GAS), *S. agalactiae* (GBS), and *S. aureus*, respectively.

Further, the present invention is directed to polynucleotides derived  
30 from gram positive bacteria and which are at least partially homologous to

the polynucleotides making up the gene coding for the previously disclosed Sp36 gene of *S. pneumoniae* (U.S. Application Serial No. 60/113,048).

5 The present invention is also directed to polynucleotides, and immunologically active fragments, segments, or portions, thereof, which polypeptides are encoded by the polynucleotides disclosed herein.

The present invention also relates to such polynucleotides and polypeptides in enriched, preferably isolated, or even purified, form.

10

In accordance with the present invention, the term "DNA segment" refers to a DNA polymer, in the form of a separate fragment or as a component of a larger DNA construct, which has been derived from DNA isolated at least once in substantially pure form, i.e., free of  
15 contaminating endogenous materials and in a quantity or concentration enabling identification, manipulation, and recovery of the segment and its component nucleotide sequences by standard biochemical methods, for example, using a cloning vector. Such segments are provided in the form of an open reading frame uninterrupted by internal nontranslated  
20 sequences, or introns, which are typically present in eukaryotic genes. Sequences of non-translated DNA may be present downstream from the open reading frame, where they do not interfere with manipulation or expression of the coding regions.

25 The nucleic acids and polypeptide expression products disclosed according to the present invention, as well as expression vectors containing such nucleic acids and/or such polypeptides, may be in "enriched form." As used herein, the term "enriched" means that the concentration of the material is at least about 2, 5, 10, 100, or 1000  
30 times its natural concentration (for example), advantageously 0.01%, by weight, preferably at least about 0.1% by weight. Enriched preparations

of about 0.5%, 1%, 5%, 10%, and 20% by weight are also contemplated. The sequences, constructs, vectors, clones, and other materials comprising the present invention can advantageously be in enriched or isolated form.

5

"Isolated" in the context of the present invention with respect to polypeptides (or polynucleotides) means that the material is removed from its original environment (*e.g.*, the natural environment if it is naturally occurring). For example, a naturally-occurring polynucleotide or polypeptide present in a living organism is not isolated, but the same polynucleotide or polypeptide, separated from some or all of the co-existing materials in the natural system, is isolated. Such polynucleotides could be part of a vector and/or such polynucleotides or polypeptides could be part of a composition, and still be isolated in that such vector or composition is not part of its natural environment. The polypeptides and polynucleotides of the present invention are preferably provided in an isolated form, and most preferably are purified to homogeneity.

The polynucleotides, and recombinant or immunogenic polypeptides, disclosed in accordance with the present invention may also be in "purified" form. The term "purified" does not require absolute purity; rather, it is intended as a relative definition, and can include preparations that are highly purified or preparations that are only partially purified, as those terms are understood by those of skill in the relevant art. For example, individual clones isolated from a cDNA library have been conventionally purified to electrophoretic homogeneity. Purification of starting material or natural material to at least one order of magnitude, preferably two or three orders, and more preferably four or five orders of magnitude is expressly contemplated. Furthermore, claimed polypeptides having a purity of preferably 0.001%, or at least 0.01% or 0.1%; and even 1% by weight or greater is expressly contemplated.

The term "coding region" refers to that portion of a gene which either naturally or normally codes for the expression product of that gene in its natural genomic environment, i.e., the region coding *in vivo* for the native expression product of the gene. The coding region can be from a normal, mutated or altered gene, or can even be from a DNA sequence, or gene, wholly synthesized in the laboratory using methods well known to those of skill in the art of DNA synthesis.

10 In accordance with the present invention, the term "nucleotide sequence" refers to a heteropolymer of deoxyribonucleotides. Generally, DNA segments encoding the proteins provided by this invention are assembled from cDNA fragments and short oligonucleotide linkers, or from a series of oligonucleotides, to provide a synthetic gene which is  
15 capable of being expressed in a recombinant transcriptional unit comprising regulatory elements derived from a microbial or viral operon.

The term "expression product" means that polypeptide or protein that is the natural translation product of the gene and any nucleic acid sequence coding equivalents resulting from genetic code degeneracy and  
20 thus coding for the same amino acid(s).

The term "fragment," when referring to a coding sequence, means a portion of DNA comprising less than the complete coding region whose  
25 expression product retains essentially the same biological function or activity as the expression product of the complete coding region.

The term "primer" means a short nucleic acid sequence that is paired with one strand of DNA and provides a free 3'OH end at which a  
30 DNA polymerase starts synthesis of a deoxyribonucleotide chain.

The term "promoter" means a region of DNA involved in binding of RNA polymerase to initiate transcription.

The term "open reading frame (ORF)" means a series of triplets  
5 coding for amino acids without any termination codons and is a sequence (potentially) translatable into protein.

As used herein, reference to a DNA sequence includes both single stranded and double stranded DNA. Thus, the specific sequence, unless  
10 the context indicates otherwise, refers to the single strand DNA of such sequence, the duplex of such sequence with its complement (double stranded DNA) and the complement of such sequence.

In accordance with the present invention, the term "percent identity"  
15 or "percent identical," when referring to a sequence, means that a sequence is compared to a claimed or described sequence after alignment of the sequence to be compared (the "Compared Sequence") with the described or claimed sequence (the "Reference Sequence"). The Percent Identity is then determined according to the following formula:

20

$$\text{Percent Identity} = 100 [1 - (C/R)]$$

wherein C is the number of differences between the Reference Sequence  
25 and the Compared Sequence over the length of alignment between the Reference Sequence and the Compared Sequence wherein (i) each base or amino acid in the Reference Sequence that does not have a corresponding aligned base or amino acid in the Compared Sequence and (ii) each gap in the Reference Sequence and (iii) each aligned base or amino acid in the  
30 Reference Sequence that is different from an aligned base or amino acid in the Compared Sequence, constitutes a difference; and R is the number of

bases or amino acids in the Reference Sequence over the length of the alignment with the Compared Sequence with any gap created in the Reference Sequence also being counted as a base or amino acid.

5           If an alignment exists between the Compared Sequence and the Reference Sequence for which the percent identity as calculated above is about equal to or greater than a specified minimum Percent Identity then the Compared Sequence has the specified minimum percent identity to the Reference Sequence even though alignments may exist in which the  
10   hereinabove calculated Percent Identity is less than the specified Percent Identity.

          Thus, the present invention is directed to novel, isolated polypeptides comprising an amino acid sequence at least 75% identical to a sequence in  
15   SEQ ID NO: 2, 4 or 6, preferably polypeptides at least 90% identical thereto, more preferably 95% identical to the sequence of SEQ ID NO: 2 or 4, and most preferably having the sequence of either SEQ ID NO: 2 or 4.

          The isolated polypeptides of the present invention may be found in a  
20   wide variety of microorganisms, but will commonly be found in an organism selected from the group consisting of group A streptococci, group B streptococci, and *Staphylococcus aureus*, and wherein the group A streptococcal organism is *Streptococcus pyogenes* and the group B streptococcal organism is *Streptococcus agalactiae*. Also, polypeptides of  
25   the invention include, but are in no way limited to, isolated polypeptides having a sequence at least 25% identical to the amino acid sequence of the Sp36 protein of *Streptococcus pneumoniae*.

          The present invention further relates to immunogenically active  
30   fragments of the isolated polypeptides disclosed herein.

The terms "fragment," "derivative" and "analog" when referring to the polypeptides disclosed herein means a polypeptide which retains essentially the same biological function or activity as such polypeptide. Thus, an analog includes a proprotein, or preprotein, which can be  
5 activated by cleavage of the proprotein portion to produce an active mature polypeptide. Such fragments, derivatives and analogs must have sufficient similarity to the polypeptide of SEQ ID NO:2, 4 or 6 so that immunogenic activity of the native polypeptide is retained.

The polypeptide of the present invention may be a recombinant  
10 polypeptide, a natural polypeptide or a synthetic polypeptide, preferably a recombinant polypeptide.

The fragment, derivative or analog of the polypeptide of SEQ ID NO:2, 4, or 6 may be (i) one in which one or more of the amino acid  
15 residues are substituted with a conserved or non-conserved amino acid residue (preferably a conserved amino acid residue) and such substituted amino acid residue may or may not be one encoded by the genetic code, or (ii) one in which one or more of the amino acid residues includes a substituent group, or (iii) one in which the mature polypeptide is fused with  
20 another compound, such as a compound to increase the half-life of the polypeptide (for example, polyethylene glycol), or (iv) one in which the additional amino acids are fused to the mature polypeptide, such as a leader or secretory sequence or a sequence which is employed for purification of the mature polypeptide or a proprotein sequence. Such fragments,  
25 derivatives and analogs are deemed to be within the scope of those skilled in the art from the teachings herein.

As known in the art "similarity" between two polypeptides is determined by comparing the amino acid sequence and its conserved amino  
30 acid substitutes of one polypeptide to the sequence of a second polypeptide.

Fragments or portions of the polypeptides of the present invention may be employed for producing the corresponding full-length polypeptide by peptide synthesis; therefore, the fragments may be employed as  
5 intermediates for producing the full-length polypeptides. Fragments or portions of the polynucleotides of the present invention may be used to synthesize full-length polynucleotides of the present invention.

As used herein with reference to polypeptides, the terms "portion,"  
10 "segment," and "fragment," refer also to a continuous sequence of residues, such as amino acid residues, which sequence forms a subset of a larger sequence. For example, if a polypeptide were subjected to treatment with any of the common endopeptidases, such as trypsin, chymotrypsin, or papain, the oligopeptides resulting from such treatment would represent  
15 portions, segments or fragments of the starting polypeptide.

The present invention is also directed to isolated polynucleotides whose sequences contain coding regions encoding the polypeptides of the present invention, preferably the polypeptides of SEQ ID NO: 2, 4, and 6  
20 and most preferably will be the isolated polynucleotides comprising the sequences of SEQ ID NOS: 1, 3, and 5.

The present invention is also directed to fragments or portions of such sequences which contain at least 15 bases, preferably at least 30  
25 bases, more preferably at least 50 bases and most preferably at least 80 bases, and to those sequences which are at least 60%, preferably at least 80%, and most preferably at least 95%, especially 98%, identical thereto, and to DNA (or RNA) sequences encoding the same polypeptide as the sequences of SEQ ID NOS: 2, 4, and 6 including fragments and  
30 portions thereof and, when derived from natural sources, includes alleles thereof.

Yet another aspect of the present invention is directed to an isolated DNA (or RNA) sequence or molecule comprising at least the coding region of a bacterial gene (or a DNA sequence encoding the same polypeptide as such coding region), in particular an expressed bacterial gene, which bacterial gene comprises a DNA sequence homologous with, or contributing to, the sequence depicted in SEQ ID NOS: 1, 3, and 5 or one at least 60%, preferably at least 80%, and most preferably at least 95%, especially 98%, identical thereto, including 100% identity, as well as fragments or portions of the coding region which encode a polypeptide having a similar function to the polypeptide encoded by said coding region. Thus, the isolated DNA (or RNA) sequence may include only the coding region of the expressed gene (or fragment or portion thereof as hereinabove indicated) or may further include all or a portion of the non-coding DNA (or RNA) of the expressed bacterial gene.

In general, sequences homologous with and contributing to the sequences of SEQ ID NOS: 1, 3, and 5 (or one at least 60%, preferably at least 80%, and most preferably at least 95% identical or homologous thereto) are from the coding region of a bacterial gene.

The polynucleotides according to the present invention may also occur in the form of mixtures of polynucleotides hybridizable to some extent with the gene sequences containing any of the nucleotide sequences of SEQ ID NOS: 1, 3, and 5, including any and all fragments thereof, and which polynucleotide mixtures may be composed of any number of such polynucleotides, or fragments thereof, including mixtures having at least 10, perhaps at least 30 such sequences, or fragments thereof.

Fragments of the full length polynucleotide of the present invention may be used as hybridization probes for a DNA library to isolate the full

length DNA and to isolate other DNAs which have a high sequence similarity to the gene or similar biological activity. Probes of this type preferably have at least 15 bases, may have at least 30 bases and even 50 or more bases. The probe may also be used to identify a DNA clone  
5 corresponding to a full length transcript and a genomic clone or clones that contain the complete gene including regulatory and promotor regions. An example of a screen comprises isolating the coding region of the gene by using the known DNA sequence to synthesize an oligonucleotide probe. Labeled oligonucleotides having a sequence complementary to that of the  
10 gene of the present invention are used to screen a library of DNA or mRNA to determine which members of the library the probe hybridizes to.

The present invention is also directed to vectors comprising the polynucleotides disclosed herein, as well as to genetically engineered cells  
15 comprising such vectors and/or polynucleotides. Thus, the present invention also relates to vectors which include polynucleotides of the present invention, host cells which are genetically engineered with vectors of the invention and the production of polypeptides of the invention by recombinant techniques.

20

Host cells are genetically engineered (transduced or transformed or transfected) with the vectors of this invention which may be, for example, a cloning vector or an expression vector. The vector may be, for example, in the form of a plasmid, a viral particle, a phage, etc. The engineered host  
25 cells can be cultured in conventional nutrient media modified as appropriate for activating promoters, selecting transformants or amplifying the genes of the present invention. The culture conditions, such as temperature, pH and the like, are those previously used with the host cell selected for expression, and will be apparent to the ordinarily skilled artisan.

30

The polynucleotides of the present invention may be employed for producing polypeptides by recombinant techniques. Thus, for example, the polynucleotide may be included in any one of a variety of expression vectors for expressing a polypeptide. Such vectors include chromosomal, nonchromosomal and synthetic DNA sequences, e.g., derivatives of SV40; bacterial plasmids; phage DNA; baculovirus; yeast plasmids; vectors derived from combinations of plasmids and phage DNA, viral DNA such as vaccinia, adenovirus, fowl pox virus, and pseudorabies. However, any other vector may be used as long as it is replicable and viable in the host.

10

The appropriate DNA sequence may be inserted into the vector by a variety of procedures. In general, the DNA sequence is inserted into an appropriate restriction endonuclease site(s) by procedures known in the art. Such procedures and others are deemed to be within the scope of those skilled in the art.

15

The DNA sequence in the expression vector is operatively linked to an appropriate expression control sequence(s) (promoter) to direct mRNA synthesis. As representative examples of such promoters, there may be mentioned: LTR or SV40 promoter, the E. coli. lac or trp, the phage lambda P<sub>L</sub> promoter and other promoters known to control expression of genes in prokaryotic or eukaryotic cells or their viruses. The expression vector also contains a ribosome binding site for translation initiation and a transcription terminator. The vector may also include appropriate sequences for amplifying expression.

25

In addition, the expression vectors preferably contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells such as dihydrofolate reductase or neomycin resistance for eukaryotic cell culture, or such as tetracycline or ampicillin resistance in E. coli.

30

The vector containing the appropriate DNA sequence as hereinabove described, as well as an appropriate promoter or control sequence, may be employed to transform an appropriate host to permit the host to express the  
5 protein.

As representative examples of appropriate hosts, there may be mentioned: bacterial cells, such as E. coli, Streptomyces, Salmonella typhimurium; fungal cells, such as yeast; insect cells such as Drosophila S2  
10 and Spodoptera Sf9; animal cells such as CHO, COS or Bowes melanoma; adenoviruses; plant cells, etc. The selection of an appropriate host is deemed to be within the scope of those skilled in the art from the teachings herein.

15 More particularly, the present invention also includes recombinant constructs comprising one or more of the sequences as broadly described above. The constructs comprise a vector, such as a plasmid or viral vector, into which a sequence of the invention has been inserted, in a forward or reverse orientation. In a preferred aspect of this embodiment, the construct  
20 further comprises regulatory sequences, including, for example, a promoter, operably linked to the sequence. Large numbers of suitable vectors and promoters are known to those of skill in the art, and are commercially available. The following vectors are provided by way of example; Bacterial: pQE70, pQE60, pQE-9 (Qiagen), pBS, pD10, phagescript, phiX174,  
25 pBluescript SK, pBSKS, pNH8A, pNH16a, pNH18A, pNH46A (Stratagene); pTRC99a, pKK223-3, pKK233-3, pDR540, pRIT5 (Pharmacia); Eukaryotic: pWLNEO, pSV2CAT, pOG44, pXT1, pSG (Stratagene) pSVK3, pBPV, pMSG, pSVL (Pharmacia). However, any other plasmid or vector may be used as long as they are replicable and viable in the host.

30

Promoter regions can be selected from any desired gene using CAT (chloramphenicol transferase) vectors or other vectors with selectable markers. Two appropriate vectors are pKK232-8 and pCM7. Particular named bacterial promoters include lacI, lacZ, T3, T7, gpt, lambda P<sub>R</sub>, P<sub>L</sub> and trp. Eukaryotic promoters include CMV immediate early, HSV thymidine kinase, early and late SV40, LTRs from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art.

10 In a further embodiment, the present invention relates to host cells containing the above-described constructs. The host cell can be a higher eukaryotic cell, such as a mammalian cell, or a lower eukaryotic cell, such as a yeast cell, or the host cell can be a prokaryotic cell, such as a bacterial cell. Introduction of the construct into the host cell can be effected by  
15 calcium phosphate transfection, DEAE-Dextran mediated transfection, or electroporation (Davis, L., Digner, M., Battey, I., Basic Methods in Molecular Biology, (1986)).

The constructs in host cells can be used in a conventional manner to  
20 produce the gene product encoded by the recombinant sequence. Alternatively, the polypeptides of the invention can be synthetically produced by conventional peptide synthesizers.

Mature proteins can be expressed in mammalian cells, yeast,  
25 bacteria, or other cells under the control of appropriate promoters. Cell-free translation systems can also be employed to produce such proteins using RNAs derived from the DNA constructs of the present invention. Appropriate cloning and expression vectors for use with prokaryotic and eukaryotic hosts are described by Sambrook, et al., Molecular Cloning: A  
30 Laboratory Manual, Second Edition, Cold Spring Harbor, N.Y., (1989), the disclosure of which is hereby incorporated by reference.

Transcription of the DNA encoding the polypeptides of the present invention by higher eukaryotes is increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp that act on a promoter to increase its transcription. Examples include the SV40 enhancer on the late side of the replication origin bp 100 to 270, a cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers.

10

Generally, recombinant expression vectors will include origins of replication and selectable markers permitting transformation of the host cell, e.g., the ampicillin resistance gene of E. coli and S. cerevisiae Trp1 gene, and a promoter derived from a highly-expressed gene to direct transcription of a downstream structural sequence. Such promoters can be derived from operons encoding glycolytic enzymes such as 3-phosphoglycerate kinase (PGK),  $\alpha$ -factor, acid phosphatase, or heat shock proteins, among others. The heterologous structural sequence is assembled in appropriate phase with translation initiation and termination sequences, and preferably, a leader sequence capable of directing secretion of translated protein into the periplasmic space or extracellular medium. Optionally, the heterologous sequence can encode a fusion protein including an N-terminal identification peptide imparting desired characteristics, e.g., stabilization or simplified purification of expressed recombinant product.

20  
25

Useful expression vectors for bacterial use are constructed by inserting a structural DNA sequence encoding a desired protein together with suitable translation initiation and termination signals in operable reading phase with a functional promoter. The vector will comprise one or more phenotypic selectable markers and an origin of replication to ensure maintenance of the vector and to, if desirable, provide amplification within

30

the host. Suitable prokaryotic hosts for transformation include E. coli, Bacillus subtilis, Salmonella typhimurium and various species within the genera *Pseudomonas*, *Streptomyces*, and *Staphylococcus*, although others may also be employed as a matter of choice.

5

As a representative but nonlimiting example, useful expression vectors for bacterial use can comprise a selectable marker and bacterial origin of replication derived from commercially available plasmids comprising genetic elements of the well known cloning vector pBR322 (ATCC 37017).

10 Such commercial vectors include, for example, pKK223-3 (Pharmacia Fine Chemicals, Uppsala, Sweden) and GEM1 (Promega Biotec, Madison, WI, USA). These pBR322 "backbone" sections are combined with an appropriate promoter and the structural sequence to be expressed.

15 Following transformation of a suitable host strain and growth of the host strain to an appropriate cell density, the selected promoter is induced by appropriate means (e.g., temperature shift or chemical induction) and cells are cultured for an additional period.

20 Cells are typically harvested by centrifugation, disrupted by physical or chemical means, and the resulting crude extract retained for further purification.

Microbial cells employed in expression of proteins can be disrupted  
25 by any convenient method, including freeze-thaw cycling, sonication, mechanical disruption, or use of cell lysing agents, such methods are well known to those skilled in the art.

Various mammalian cell culture systems can also be employed to  
30 express recombinant protein. Examples of mammalian expression systems include the COS-7 lines of monkey kidney fibroblasts, described by

Gluzman, Cell, 23:175 (1981), and other cell lines capable of expressing a compatible vector, for example, the C127, 3T3, CHO, HeLa and BHK cell lines. Mammalian expression vectors will comprise an origin of replication, a suitable promoter and enhancer, and also any necessary ribosome binding sites, polyadenylation site, splice donor and acceptor sites, transcriptional termination sequences, and 5' flanking nontranscribed sequences. DNA sequences derived from the SV40 splice, and polyadenylation sites may be used to provide the required nontranscribed genetic elements.

10 The polypeptide can be recovered and purified from recombinant cell cultures by methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxylapatite chromatography and lectin  
15 chromatography. Protein refolding steps can be used, as necessary, in completing configuration of the mature protein. Finally, high performance liquid chromatography (HPLC) can be employed for final purification steps.

The polypeptides of the present invention may be a naturally purified  
20 product, or a product of chemical synthetic procedures, or produced by recombinant techniques from a prokaryotic or eukaryotic host (for example, by bacterial, yeast, higher plant, insect and mammalian cells in culture). Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated or may be  
25 non-glycosylated. Polypeptides of the invention may also include an initial methionine amino acid residue.

The polypeptides of the present invention, when utilized for clinically related purposes, may also be suspended in a pharmacologically  
30 acceptable diluent or excipient to facilitate such uses, which will include

use as a vaccine for the purpose of preventing a wide variety of streptococcal and staphylococcal infections.

In accordance with another aspect of the present invention, there is provided a vaccine that includes at least one polypeptide that is at least 75% identical, preferably at least 90% identical and most preferably 95% identical, to a polypeptide sequence comprising the sequence of SEQ ID NO: 2, 4, or 6. Such variations in homology for putative vaccines are well known in the art (See, for example, Hanson et al., "Active and Passive Immunity Against *Borrelia burgdorferi* Decorin Binding Protein A (DbpA)," Infection and Immunity, (May) 1998, p. 2143 – 2153; Roberts et al., "Heterogeneity Among Genes Including Decorin Binding Proteins A and B of *Borrelia burgdorferi sensu lato*," Infection and Immunity, (Nov) 1998, p. 5275-5285). Such observations would similarly apply to portions, segments or fragments of the polypeptides disclosed herein.

Such segments find a multitude of uses. For example, such segments of the polypeptides according to the present invention find use as intermediates in the synthesis of higher molecular weight structures also within the present invention.

The term "active fragment" means a fragment that generates an immune response (i.e., has immunogenic activity) when administered, alone or optionally with a suitable adjuvant, to an animal, such as a mammal, for example, a rabbit or a mouse, and also including a human.

In accordance with a further aspect of the invention, a vaccine of the type hereinabove described is administered for the purpose of preventing or treating infection caused by streptococci and staphylococci as well as many related organisms.

A vaccine in accordance with the present invention may include one or more of the hereinabove described polypeptides or active fragments thereof. When employing more than one polypeptide or active fragment, such as two or more polypeptides and/or active fragments may be used as  
5 a physical mixture or as a fusion of two or more polypeptides or active fragments. The fusion fragment or fusion polypeptide may be produced, for example, by recombinant techniques or by the use of appropriate linkers for fusing previously prepared polypeptides or active fragments.

10 In many cases, the variation in the polypeptide or active fragment is a conservative amino acid substitution, although other substitutions are within the scope of the invention.

In accordance with the present invention, a polypeptide variant  
15 includes variants in which one or more amino acids are substituted and/or deleted and/or inserted.

In another aspect, the invention relates to passive immunity vaccines formulated from antibodies against a polypeptide or active fragment of a  
20 polypeptide of the present invention. Such passive immunity vaccines can be utilized to prevent and/or treat streptococcal and staphylococcal infections in patients. In this manner, according to a further aspect of the invention, a vaccine can be produced from a synthetic or recombinant polypeptide of the present invention or an antibody against such  
25 polypeptide.

Still another aspect the present invention relates to a method of using one or more antibodies (monoclonal, polyclonal or sera) to the polypeptides of the invention as described above for the prophylaxis and/or treatment of  
30 diseases that are caused by streptococcal and staphylococcal bacteria. In particular, the invention relates to a method for the prophylaxis and/or

treatment of infectious diseases that are caused by streptococci and staphylococci. In a still further preferred aspect, the invention relates to a method for the prophylaxis and/or treatment of such diseases as necrotizing fasciitis, scarlet fever, sepsis and many diseases of newborns, in humans  
5 by utilizing a vaccine of the present invention.

Generally, vaccines are prepared as injectables, in the form of aqueous solutions or suspensions. Vaccines in an oil base are also well known such as for inhaling. Solid forms which are dissolved or suspended  
10 prior to use may also be formulated. Pharmaceutical carriers, diluents and excipients are generally added that are compatible with the active ingredients and acceptable for pharmaceutical use. Examples of such carriers include, but are not limited to, water, saline solutions, dextrose, or glycerol. Combinations of carriers may also be used.

15

Vaccine compositions may further incorporate additional substances to stabilize pH, or to function as adjuvants, wetting agents, or emulsifying agents, which can serve to improve the effectiveness of the vaccine.

20 Vaccines are generally formulated for parenteral administration and are injected either subcutaneously or intramuscularly. Such vaccines can also be formulated as suppositories or for oral administration, using methods known in the art, or for administration through nasal or respiratory routes.

25

The amount of vaccine sufficient to confer immunity to pathogenic bacteria is determined by methods well known to those skilled in the art. This quantity will be determined based upon the characteristics of the vaccine recipient and the level of immunity required. Typically, the amount  
30 of vaccine to be administered will be determined based upon the judgment of a skilled physician. Where vaccines are administered by subcutaneous or

intramuscular injection, a range of 0.5 to 500 µg purified protein may be given.

5 The present invention is also directed to a vaccine in which a polypeptide or active fragment of the present invention is delivered or administered in the form of a polynucleotide encoding the polypeptide or active fragment, whereby the polypeptide or active fragment is produced *in vivo*. The polynucleotide may be included in a suitable expression vector and combined with a pharmaceutically acceptable carrier.

10

Thus, the present invention expressly contemplates a vaccine composition comprising any of the polypeptides disclosed herein, said polypeptide being present in an amount effective to produce an immune response, and wherein said polypeptide is suspended in a pharmacologically acceptable carrier, diluent or excipient.

15

The vaccine compositions of the present invention may also comprise live vaccines, containing such organisms as *Streptococcus gordoniae* and *Salmonella typhi*, wherein said organisms contain recombinant polypeptides as disclosed herein.

20

In addition, the polypeptides of the present invention can be used as immunogens to stimulate the production of antibodies for use in passive immunotherapy, for use as diagnostic reagents, and for use as reagents in other processes such as affinity chromatography.

25

Thus, the present invention is also directed to methods for the prevention of a wide variety of diseases caused by streptococcal and staphylococcal organisms, said methods involving the administering of vaccines disclosed herein to animals at risk of such diseases, especially where said animals are humans.

30

In addition, the invention disclosed herein is also directed to a means of treating animals, especially humans, afflicted with a disease caused by the organisms from which the isolated polypeptides of the invention are derived, such methods including, but not being limited to, administering to an animal, especially a human, afflicted with such a disease of a therapeutically effective amount of an antibody, or mixture of antibodies, against the polypeptides disclosed herein.

Antibodies specific for the polypeptides disclosed herein may be either polyclonal or monoclonal and may even be in the form of antisera. When such antibodies are monoclonal in nature, they may be produced by conventional methods of preparing monoclonal antibodies, such as from conventional hybridoma cells, and may also be produced by genetically engineered cells transformed with vectors containing genes specifically coding for the different heavy and light chains of antibody molecules having an arrangement of variable regions specifically complementary to one or more of the polypeptides of the invention. Such recombinantly produced antibodies may be in the form of either dimers or tetramers, depending on the type of cellular expression system utilized therefor.

The invention will now be further described in more detail in the following non-limiting examples and it will be appreciated that additional and different embodiments of the teachings of the present invention will doubtless suggest themselves to those of skill in the art and such other embodiments are considered to have been inferred from the disclosure herein.

## Example 1

**Southern Blot Analysis of Chromosomal DNA Using Probes Specific for the Sp36 Gene of *Streptococcus pneumoniae***

5

Genomic DNA was isolated from *Staphylococcus aureus*, *Streptococcus pyogenes* (group A), and *Streptococcus agalactiae* (group B) after overnight growth of the bacteria. The DNA was digested to completion by overnight incubation with restriction enzymes (*Bam*HI and *Pvu*II), and then DNA fragments were resolved by size by agarose gel electrophoresis before transfer to a nylon membrane. The membrane was then probed with DNA encoding the entire Sp36 open reading frame that had been fluorescein-labeled with random primers using a kit from Amersham Pharmacia Biotech Inc. The hybridization and washes were carried out under low stringency conditions (i.e., 45°C, 5xSSC hybridization; 45°C, 1xSSC for 1<sup>st</sup> wash; 45°C, 0.5xSSC for 2<sup>nd</sup> wash). Here, SSC is composed of 150 mM NaCl and 15 mM sodium citrate, pH 7.0 and all washes are 50 mL each.

After hybridization and washing was complete, the bound, fluorescein-labeled probe was detected using an anti-fluorescein antibody as per the manufacturer's instructions with the kit. Similarly digested DNA from *Streptococcus pneumoniae* strain SJ2 (serotype 6B) was used as a positive control. Fluorescein-labeled bacteriophage lambda DNA digested with the restriction nuclease *Hind*III was used as a size marker.

The Sp36 probe hybridized with a single fragment in the digested *S. aureus* DNA (~4.5 kb *Bam*HI fragment, ~5 kb *Pvu*II fragment) and with 2 major fragments in a *Pvu*II digest of serotype M1 of the group A streptococci genomic DNA (~4.0 kb, and ~4.2 kb ).

## Example 2

### BLAST Analysis Using Sp36 Predicted Amino Acid Sequence

5           Sequence comparisons of the Sp36 encoded protein sequence against the publicly available GenBank sequence database (including the unfinished microbial database (<http://www.ncbi.nlm.nih.gov/BLAST/unfinishedgenome.html>)) revealed two highly homologous amino acid sequences. One of these was a predicted amino acid sequence from the *S. pyogenes* genome. This predicted polypeptide comprised 825 amino acid residues (MW = 92,616 Da) that was 25.1% identical to the Sp36 amino acid sequence from pneumococcus serotype 4 but maintained the 5 histidine triads (underlined in Figure 5(a) - SEQ ID NO: 2). The second polypeptide encoded within the *S. pyogenes* database contained several errors that were corrected by our sequencing of this region of the genome. The DNA fragment obtained encoded a protein of 792 amino acids (MW = 87,457 Da) that was 12.6% identical to the pneumococcal sequence and 12.5% identical to the first *S. pyogenes* polypeptide. This predicted amino acid sequence contained four histidine triad motifs (underlined in Fig. 5(b) - SEQ ID NO.: 4). The third polypeptide sequence obtained was one already in the database (Accession No. AF062533) and identified only as an unknown gene downstream from a gene identified as *lmb* in *S. galactiae*. This 822 amino acid protein thus has a predicted molecular weight of 92,353 Da and maintains the 5 histidine triad motifs (underlined in Figure 5(C) - SEQ ID NO: 6). This second polypeptide shows 25.6% sequence identity to Sp36 of pneumococcus type 4 and 97.7% and 11.6% identity to the two group A homologs, respectively.

### Example 3

#### Southern Blot Analysis Using a group A Streptococcal Sp36 Homolog Probe

5 Southern blot analysis was performed with a fluorescein-labeled DNA fragment as probe, which encoding a group A streptococcal Sp36 homolog cloned from an M1 serotype of the group A streptococcal genome. This fragment was then used to probe genomic DNA from an M6 serotype of the group A streptococcal genome, as well as serotype 1a and serotype 3 of  
10 the group B streptococcal genome, and strain SJ2 (serotype 6B) of pneumococcus. In all cases, a single band was obtained in DNA digested with *Bam*HI when hybridization was carried out under low stringency conditions (as described above). A band of about 20 kb was visualized in group A streptococcal DNA, about 4.5 kb was obtained for group B  
15 streptococcal DNA, and a band of about 4kb was seen for pneumococcus.

### Example 4

20

#### Western Blot Analysis of Reactivity of group B Streptococcal Homolog With Anti-Pneumococcal Sp36 Antiserum

To determine whether antiserum raised against recombinant Sp36  
25 from *S. pneumoniae* would recognize the recombinant Sp36 homolog encoded by group B streptococcal organisms, a western blot was performed. One hundred nanograms (100 ng) of recombinant Sp36 polypeptide cloned from either *S. pneumoniae* serotype 4, or of the Sp36 homolog cloned from group B streptococcal organisms, or from an unrelated  
30 recombinant protein control expressed and purified in the same way, were subjected to SDS-PAGE containing 12% acrylamide. A cell lysate of

pneumococcal strain SJ2 (serotype 6B) was also included on the gel. After electrophoresis, the separated proteins were transferred to a nitrocellulose membrane and probed with rabbit polyclonal antiserum raised against the recombinant pneumococcal protein. Bound antibodies were detected  
5 chemiluminescently with a goat anti-rabbit IgG antibody conjugated to horseradish peroxidase using the substrate ECL (from Amersham). The results demonstrate that antiserum raised against the pneumococcal Sp36 protein cross-react with the Sp36 homolog identified from the group B streptococci and thereby indicating conservation of epitopes between the  
10 proteins. The group B streptococcal homolog is also approximately the same size as the protein detected in *S. pneumoniae* lysates. Because the group A and B homologs are highly homologous, if not identical, such antiserum would also likely cross-react with the group A streptococcal protein.

15

#### Example 5

##### Alignment of Predicted Amino Acid Sequences of the Sp36 Homologs from group A and B Streptococci With Pneumococcal Sp36

20

The predicted amino acid sequences from the Sp36 genes from group A and group B streptococci and *S. pneumoniae* were aligned using the Clustal algorithm in a DNASTar Computer package (DNASTar, Inc., Madison, WI). Amino acids that match those encoded by the pneumococcal  
25 gene are boxed in Figure 2 (showing the results of the alignment). Gaps introduced in the sequence by the alignment process are indicated by dashed lines.

30

#### Example 6

### Percent Sequence Identity Between Homologs of Sp36

The Sp36 amino acid sequence from pneumococci is 25.6% identical to the predicted amino acid sequence of the homologous gene of group B streptococci and 25.1% and 12.6% identical to the deduced sequences of the two genes from group A streptococci. Furthermore, the group B homolog is 97.7% and 11.6% identical to the first (GAS36) and second (GAS36(2)) homologs from group A streptococci, respectively. These experiments indicate that homologous genes to Sp36 from pneumococcus are present in group A and group B streptococci, as well as in *Staphylococcus aureus*. The protein encoded by this gene may therefore perform a similar function in these different organisms. This suggests that a vaccine comprising one or more of these proteins may be broadly protective against these species. These results are summarized in Table 1 which shows the percent identity between the amino acid sequences of Sp36 from pneumococcus strain Norway 4 (serotype 4), group A streptococci Sp36 homolog from an M1 serotype, and group B streptococci Sp36 from strain R268.

Table 1.

	Pneumo. Sp36	GAS36	GAS36(2)	GBS36
Pneumo. Sp36	100%	25.1%	12.6%	25.6%
GAS36	---	100%		97.7%
GAS36(2)	---	---	100%	11.6%
GBS36	---	---		100%

where GAS36 = SEQ ID NO: 2

GAS36(2) = SEQ ID NO: 4

GBS36 = SEQ ID NO: 6

WHAT IS CLAIMED IS:

1. An isolated polypeptide comprising an amino acid sequence at least 75% identical to a sequence selected from the group consisting of  
5 SEQ ID NO: 2, 4 and 6.
2. The isolated polypeptide of claim 1 wherein said polypeptide is at least 90% identical to the sequence selected from the group consisting of  
10 SEQ ID NO: 2, 4, and 6.
3. The isolated polypeptide of claim 1 wherein said polypeptide is at least 95% identical to the sequence selected from the group consisting of  
15 SEQ ID NO: 2, 4, and 6.
4. The isolated polypeptide of claim 1 wherein said polypeptide has the amino acid sequence selected from the group consisting of SEQ ID NO: 2, 4 and 6.
5. The isolated polypeptide of claim 1 wherein said polypeptide is  
20 found in an organism selected from the group consisting of group A streptococci, group B streptococci, and *Staphylococcus aureus*.
6. The isolated polypeptide of claim 5 wherein the group A streptococcal organism is *Streptococcus pyogenes*.  
25
7. The isolated polypeptide of claim 5 wherein the group B streptococcal organism is *Streptococcus agalactiae*.
8. The isolated polypeptide of claim 1 wherein said polypeptide has a  
30 sequence at least 25% identical to the amino acid sequence of the Sp36 protein of *Streptococcus pneumoniae*.

9. An isolated polynucleotide comprising a sequence coding for a polypeptide selected from the group consisting of the polypeptides of claims 1, 2, 3, 4, 5, 6, 7, and 8.

5

10. The isolated polynucleotide of claim 9 wherein said polynucleotide has a nucleotide sequence selected from the group consisting of SEQ ID NO: 1, 3 and 5.

10 11. An antibody specific for a polypeptide selected from the group consisting of the polypeptides of claims 1, 2, 3, 4, 5, 6, 7, and 8.

12. The antibody of claim 11 wherein said antibody is a monoclonal antibody.

15

13. A genetically engineered cell producing the antibody of claim 12.

14. A vector comprising the polynucleotide of claim 9.

20 15. A vector comprising the polynucleotide of claim 10.

16. A genetically engineered cell expressing the polypeptide coded for by the polynucleotide of claim 9 or 10.

25 17. A composition comprising a polypeptide selected from the group consisting of the polypeptides of claims 1, 2, 3, 4, 5, 6, 7, and 8, said polypeptide being suspended in a pharmacologically acceptable diluent or excipient.

30 18. A vaccine composition comprising a polypeptide selected from the group consisting of the polypeptide of claims 1, 2, 3, 4, 5, 6, 7, and 8,

said polypeptide being present in an amount effective to produce an immune response, and wherein said polypeptide is suspended in a pharmacologically acceptable carrier, diluent or excipient.

5           19. A vaccine comprising an immunogenically active amount of the composition of claim 17.

20. A method of vaccinating an animal against infection by a bacterial organism selected from the group consisting of streptococcal  
10 bacteria and staphylococcal bacteria comprising administering to said animal an immunologically effective amount of the vaccine of claim 19.

21. The method of claim 20 wherein said animal is a human.

15           22. A method of treating a disease comprising administering to an animal afflicted therewith of a therapeutically effective amount of an antibody of claim 12 wherein said antibody is suspended in a pharmacologically acceptable carrier, diluent or excipient.

20           23. The method of claim 22 wherein said animal is a human.

24. The method of claim 22 wherein said disease is caused by an organism selected from the group consisting of group A streptococci, group B streptococci, and *Staphylococcus aureus*.

25

30

**Figure 1**

Figure 2(a)

		10	20	30																											
1	V	K	K	T	Y	G	Y	-	I	G	S	V	A	A	I	L	L	A	T	H	I	G	S	Y	Q	L	G	K	H	H	Gas36.pro
1	V	K	K	T	Y	G	Y	-	I	G	S	V	A	A	I	L	L	A	T	H	I	G	S	Y	Q	L	G	K	H	H	Gbs36.PRO
1	H	K	I	N	K	K	Y	L	V	G	S	A	A	A	L	L	S	-	-	V	C	S	Y	E	L	G	L	Y	Q	Pneumo Sp36.PRO	
		40	50	60																											
30	M	G	S	A	T	K	D	N	Q	I	A	Y	I	D	D	S	K	G	K	A	K	A	P	K	T	N	K	T	M	D	Gas36.pro
30	M	G	L	A	T	K	D	N	Q	I	A	Y	I	D	D	S	K	G	K	V	K	A	P	K	T	N	K	T	M	D	Gbs36.PRO
29	A	R	T	V	K	E	N	N	R	V	S	Y	I	D	G	K	Q	A	T	Q	K	T	-	-	E	N	L	T	P	D	Pneumo Sp36.PRO
		70	80	90																											
60	Q	I	S	A	E	E	G	I	S	A	E	Q	I	V	V	K	I	T	D	Q	G	Y	V	T	S	H	G	D	H	Y	Gas36.pro
60	Q	I	S	A	E	E	G	I	S	A	E	Q	I	V	V	K	I	T	D	Q	G	Y	V	T	S	H	G	D	H	Y	Gbs36.PRO
57	E	V	S	K	R	E	G	I	N	A	E	Q	I	V	I	K	I	T	D	Q	G	Y	V	T	S	H	G	D	H	Y	Pneumo Sp36.PRO
		100	110	120																											
90	H	F	Y	N	G	K	V	P	Y	D	A	I	S	E	E	L	L	M	T	D	P	N	Y	R	F	K	Q	S	D	Gas36.pro	
90	H	F	Y	N	G	K	V	P	Y	D	A	I	S	E	E	L	L	M	T	D	P	N	Y	H	F	K	Q	S	D	Gbs36.PRO	
87	H	Y	Y	N	G	K	V	P	Y	D	A	I	S	E	E	L	L	M	K	D	P	N	Y	K	L	K	D	E	D	Pneumo Sp36.PRO	
		130	140	150																											
120	V	I	N	E	I	L	D	G	Y	V	I	K	V	N	G	N	Y	Y	V	Y	L	K	P	G	S	K	R	K	N	I	Gas36.pro
120	V	I	N	E	I	L	D	G	Y	V	I	K	V	N	G	N	Y	Y	V	Y	L	K	P	G	S	K	R	K	N	I	Gbs36.PRO
117	I	V	N	E	V	K	G	G	Y	V	I	K	V	D	G	K	Y	Y	V	Y	L	K	D	A	A	H	A	D	N	V	Pneumo Sp36.PRO
		160	170	180																											
150	R	T	K	Q	Q	I	A	E	Q	V	A	K	G	T	K	E	A	K	E	K	G	L	A	Q	V	A	H	L	S	K	Gas36.pro
150	R	T	K	Q	Q	I	A	E	Q	V	A	K	G	T	K	E	A	K	E	K	G	L	A	Q	V	A	H	L	S	K	Gbs36.PRO
147	R	T	K	E	E	I	N	R	Q	-	K	Q	E	H	S	D	H	R	E	G	G	T	P	R	-	-	-	-	-	-	Pneumo Sp36.PRO
		190	200	210																											
180	E	E	V	A	A	V	N	E	A	K	R	Q	G	R	Y	T	T	D	D	G	Y	I	F	S	P	T	D	I	I	D	Gas36.pro
180	E	E	V	A	A	V	N	E	A	K	R	Q	G	R	Y	T	T	D	D	G	Y	I	F	S	P	T	D	I	I	D	Gbs36.PRO
170	-	N	D	G	A	V	A	L	A	R	S	Q	G	R	Y	T	T	D	D	G	Y	I	F	N	A	S	D	I	I	E	Pneumo Sp36.PRO
		220	230	240																											
210	D	L	G	D	A	Y	L	V	P	H	G	N	H	Y	H	Y	I	P	K	K	D	L	S	P	S	E	L	A	A	A	Gas36.pro
210	D	L	G	D	A	Y	L	V	P	H	G	N	H	Y	H	Y	I	P	K	K	D	L	S	P	S	E	L	A	A	A	Gbs36.PRO
199	D	T	G	D	A	Y	I	V	P	H	G	D	H	Y	H	Y	I	P	K	N	E	L	S	A	S	E	L	A	A	A	Pneumo Sp36.PRO
		250	260	270																											
240	Q	A	Y	W	S	O	K	Q	G	R	G	A	R	P	S	D	Y	R	P	T	P	A	P	A	P	G	R	R	K	A	Gas36.pro
240	Q	A	Y	W	S	O	K	Q	G	R	G	A	R	P	S	D	Y	R	P	T	P	A	-	-	P	G	R	R	K	A	Gbs36.PRO
229	E	A	F	L	S	G	R	-	G	N	L	S	N	S	R	T	Y	R	R	-	Q	N	S	D	N	T	S	R	T	N	Pneumo Sp36.PRO
		280	290	300																											
270	P	I	P	D	V	T	P	N	P	G	Q	G	H	O	P	D	N	G	G	Y	H	P	A	P	P	R	P	N	D	A	Gas36.pro
268	P	I	P	D	V	T	P	N	P	G	Q	G	H	O	P	D	N	G	G	Y	H	P	A	P	P	R	P	N	D	A	Gbs36.PRO
257	W	V	P	S	V	S	-	-	-	-	-	-	-	-	-	-	-	-	-	-	N	P	G	T	T	N	T	N	T	S	Pneumo Sp36.PRO
		310	320	330																											
300	S	O	N	K	H	O	R	D	E	F	K	G	K	T	F	K	E	L	L	D	Q	L	H	R	L	D	L	K	Y	R	Gas36.pro
298	S	O	N	K	H	O	R	D	E	F	K	G	K	T	F	K	E	L	L	D	Q	L	H	R	L	D	L	K	Y	R	Gbs36.PRO
273	N	N	S	N	T	N	S	O	A	S	O	S	N	D	I	D	S	L	K	Q	L	Y	K	L	P	L	S	Q	R	Pneumo Sp36.PRO	

Figure 2(b)

330	H V E E D G L I F E P T Q V I K S N A F G Y V V P H G D H Y	Gas36.pro
328	H V E E D G L I F E P T Q V I K S N A F G Y V V P H G D H Y	Gbs36.PRO
303	H V E S D G L V F D P A Q I T S R T A R G V A V P H G D H Y	Pneumo Sp36.PRO
360	H I I P R S Q L S P L E M E L A D - - - R Y L A G O - - -	Gas36.pro
358	H I I P R S Q L S P L E M E L A D - - - R Y L A G O - - -	Gbs36.PRO
333	H F I P Y S Q M S E L E E R I A R I I P L R Y R S N H W V P	Pneumo Sp36.PRO
383	T E D D D S G S D H S K P S D K E V T H T F L G H R I K A Y	Gas36.pro
381	T D D N D S G S D H S K P S D K E V T H T F L G H R I K A Y	Gbs36.PRO
363	D S R P E O P S P O P T P E P S P G P O P A P N L K I D S N	Pneumo Sp36.PRO
413	G K G L D G K P Y D T S D A Y V F S K E S I I H S - V D K S G	Gas36.pro
411	G K G L D G K P Y D T S D A Y V F S K E S I I H S - V D K S G	Gbs36.PRO
393	S S L V S Q L V R K V G E G Y V F E E K G I S R Y V F A K D	Pneumo Sp36.PRO
442	V T A K H G D H F H Y - I G F G E L E O Y E L D E V A N W V	Gas36.pro
440	V T A K H G D H F H Y - I G F G E L E O Y E L D E V A N W V	Gbs36.PRO
423	L P S E T V K N L E S K L S K Q E S V S H T L T A K K E N V	Pneumo Sp36.PRO
471	K A K G Q A - - D E L A A A L D O E O G K E K P L F D T K K	Gas36.pro
469	K A K G Q A - - D E L V A A L D O E O G K E K P L F D T K K	Gbs36.PRO
453	A P R D Q E F Y D K A Y N L L T E A H - - - K A L F E N K G	Pneumo Sp36.PRO
499	V S R K V T K D G K V G Y M M P K D G K D Y F Y A R D Q L D	Gas36.pro
497	V S R K V T K D G K V G Y I M P K D G K D Y F Y A R Y Q L D	Gbs36.PRO
480	R N S D F Q A L D K L L E R L N D E S T N - - - K E K L V	Pneumo Sp36.PRO
529	L T O I A F - - - A E O E L M L K D K K H Y R Y D I V D T	Gas36.pro
527	L T O I A F - - - A E O E L M L K D K K H Y R Y D I V D T	Gbs36.PRO
506	D D L L A F L A P I T H P E R L G K P N S Q I E Y T E D E V	Pneumo Sp36.PRO
555	G I E P R L A - - - V D V S S L P M H A G N A T Y D I T G S S	Gas36.pro
553	G I E P R L A - - - V D V S S L P M H A G N A T Y D I T G S S	Gbs36.PRO
536	R I - A Q L A D K Y T T S O G Y I F D E H D I I S D E G D A	Pneumo Sp36.PRO
582	F V I P H I D H I H V V P Y S W L T - R D Q I A T I K Y V M	Gas36.pro
580	F V I P H I D H I H V V P Y S W L T - R N Q I A T I K Y V M	Gbs36.PRO
565	Y V T P H M G H S H W I G K D S L S Q K E K V A A Q A Y T K	Pneumo Sp36.PRO
611	Q H P E V R P - - - D I W S K P G H E E S G S V I P N V T	Gas36.pro
609	Q H P E V R P - - - D V W S K P G H E E S G S V I P N V T	Gbs36.PRO
595	E K G I L P P S P O A D V K A N P T G D S A A A I Y N R Y K	Pneumo Sp36.PRO

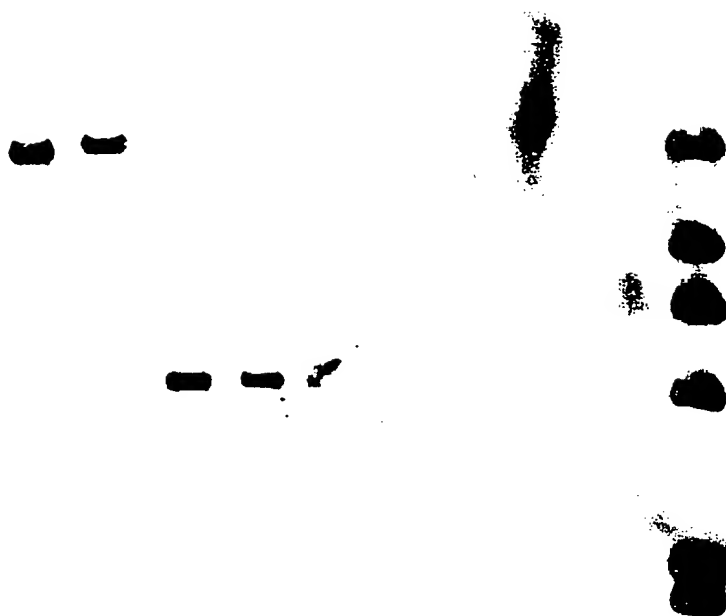
Figure 2(c)

	670	680	690	
637	P L D K R A G M P N W O - I I H S A E E V Q K A L A E G R F			Gas36.pro
635	P L D K R A G M P N W O - I I H S A E E V Q K A L A E G R F			Gbs36.PRO
625	G - E K R I P L V R L P Y M V E H T V E V K N G - - - - -			Pneumo Sp36.PRO
	700	710	720	
666	A T P D G Y I F D P R D V L A K E T F V W - K D G S F S I P			Gas36.pro
664	A A P D G Y I F D P R D V L A K E T F V W - K D G S F S I P			Gbs36.PRO
648	- - - - N L I I P H K D H Y H N I K F A V F D D H T Y K A P			Pneumo Sp36.PRO
	730	740	750	
695	R A D G S S L R T I - - - - - N K S D L S I Q A E			Gas36.pro
693	R A D G S S L R T I - - - - - N K S D L S I Q A E			Gbs36.PRO
674	- - N G Y T L E D L F A T I K Y Y V E H P D E R P H S N D G			Pneumo Sp36.PRO
	760	770	780	
714	W Q Q A Q E - L L A K K N A G D A T D T D - K P K E K Q Q A			Gas36.pro
712	W Q Q A Q E - L L A K K N A G D A T D T D - K P E E K Q Q A			Gbs36.PRO
702	W G N A S E H V L G K K D H S E D P N K N F K A D E E P V E			Pneumo Sp36.PRO
	790	800	810	
742	D K S N E N Q Q P - - - S E A S K E E E K E S D D F I D S L			Gas36.pro
740	D K S N E N Q Q P - - - S E A S K E E E K E S D D F I D S L			Gbs36.PRO
732	E T P A E P E V P O V E T E K V E A O L K E A E V L L A K V			Pneumo Sp36.PRO
	820	830	840	
769	P D Y G L D R A T L E D H I N Q L A O K A N I D P K Y L I F			Gas36.pro
766	P D Y G L D R A T L E D H I N Q L A O K A N I D P K Y L I F			Gbs36.PRO
762	T D S S L - K A N A T E T L A G L R N N I T L O - - - - I M			Pneumo Sp36.PRO
	850	860	870	
799	Q P E G V Q F Y N K N G E L V T Y - - - - D I K T L Q Q I N			Gas36.pro
796	Q P E G V Q F Y N K N G E L V T Y - - - - D I K T L Q Q I N			Gbs36.PRO
787	D N N S I H - - A E A E K L L A L L K G S N P S S V S K E K			Pneumo Sp36.PRO
825	P			Gas36.pro
822	P			Gbs36.PRO
815	I N			Pneumo Sp36.PRO

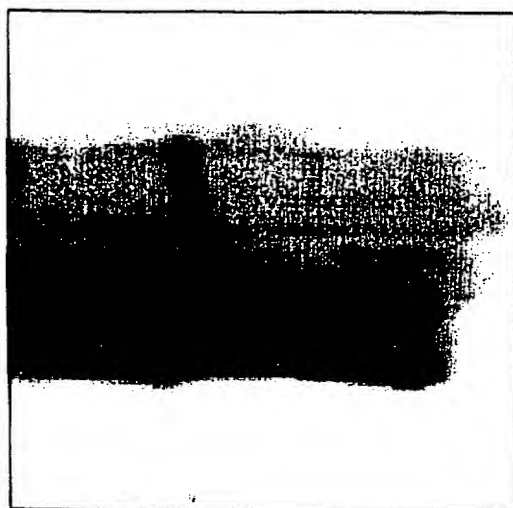
Figure 3

*S. pyogenes* M1  
*S. pyogenes* M6  
*S. agalactiae* 090R  
*S. agalactiae* 13/63  
*S. pneumoniae*

$\lambda$



**Figure 4**



Control

Sp36 GBS

Sp36

SJ2 6b

Figure 5(a)

```

MKKTYGYIGS VAAILLATHI GSYQLGKHHM GSATKDNQIA YIDDSKGKAK APKTNKTMDO ISAEEGISAE 70
QIVVKITDOG YVTSHGDHYH FYNGKVPYDA IISELLMTD PNYRFKQSDV INEILDGYVI KVNIGNYVYL 140
KPGSKRKNIR TKQIAEQVA KGTKEAKEG LAQVAHLSKE EVAAVNEAKR QGRYTTDDGY IFSPTDIIDD 210
LGDAYLVPHG NHYHYIPKD LSPSELAAQ AYWSQKQGRG ARPSDYRPTP APAPGRRKAP IPDVTPNPGQ 280
GHQPDNGGYH PAPPRPNDAS QNKHQRDEFK GKTFKELLDQ LHRDLKCYRH VEEDGLIFEP TQVIKSNAFG 350
YVVPHGHDHYH IIPRSQLSPL EMELADRYLA GQTEDDDSGS DSKPSDKEV THTFLGHRK AYKGGLDGKP 420
YDTSDAYVFS KESIHSVDS KESIAHGDHF GVTAKHGDHF HYIGFGELEQ YELDEVANWV KAKGQADELA AALDQEQKE 490
KPLFDTKKVS RKVTKDQKVG YMPKDGKDY FYARDQLDLT QIAFAEQELM LKDKKHRYD IVDTGIEPRL 560
AVDVSSLP MH AGNATYDTGS SFVIPHIDHI HVPYSWLTR DQIATIKYVM QHPEVRPDIW SKPGHEESGS 630
VIPNVTPLDK RAGMPNWI HSAEEVQKAL AEGRFATPDG YIFDPRDLA KETFWKDGGS FSIPRADGSS 700
LRTINKSDLS QAEWQQAQEL LAKKNAGDAT DTDKPKKEKQ ADKSNENQOP SEASKEEKE SDDFIDSLPD 770
YGLDRATLED HINQLAQKAN IDPKYLIFQ EGVQFYKNG ELVTYDIKTL QQINP 825

```

Figure 5(b)

```

MKTKKVILV  GLLSSQLTL  IACQSRGNGT  YPIKTKQSRK  GMTSNKIKPI  KKSCKTNKTH  KGVAGVDFPT  70
DDGFILTKDS  KILSKTDQGI  VVDHGDGSHF  IFYADLKGSP  FEYLIPKGAS  LAKPAVAQRA  ASQGTSKVAD  140
PHHHYEFNPA  DIVAEDALGY  TVRHDDHFHY  ILKSSLSGQT  QAQAKQVATR  LPQTSSSLVST  ATANGIPGLH  210
FPTSDGFQFN  GQGIVGVTKD  SILVDHDGHL  HPISFADLRQ  GGWAHVADQY  DPAKKAEPKA  ETHQTPELSE  280
REKEYQEKLA  YLAEKLGIDP  STIKRVETQD  GKLGLEYPHH  DHAHVLMUSD  IEIGKDIPDP  HAIEHARELE  350
KHKVGMDTLR  ALGFDEEVIL  DIVRTHDAPT  PFPSNEKDPN  MMKEWLATVI  KLDLGSRKOP  LQKGLSLLP  420
NLETLGIGFT  PIKDISPVLQ  FKCLKQLLMT  KTGVTDYRFL  DNMPQLEGID  ISQNNLKDIS  FLSKYKNLTL  490
VAAADNGIED  IRPLGQLPNL  KFLVLSNNKI  SDLSPLASLH  QLQELHIDNN  QITDLSPVSH  KESLTVVDLS  560
RNADVLDLATL  QAPKLETLNV  NDTKVSHLDF  LKNNPNLSSL  SINRAQLQSL  EGIEASSVIV  RVEAEGNQIK  630
SLVLKDKQGS  LTFLDVTGNQ  LTSLEGVNNF  TALDILSVSK  NQLTNVNLSK  PNKTVTNIDI  SHNNISLADL  700
KLINEQHIPEA  IAKNFPAVYE  GSMVGNGTAE  EKAAMATKAK  ESAQEASESH  DYNHNHTYED  EEGHAHEHRD  770
KDDHDEHEHD  ENEAKDEQNH  AD                                     792

```

Figure 5(c)

MKKTYGYIGS VAAILLATHI GSYQLGKHHM GLATKDNQIA YIDDSKGKVK APKTNKTMQD ISAEEGISAE 70  
 QIVVKITDQG YVTSHGHDHYH FYNGKVPYDA IISEELLMTD PNYHFKQSDV INEILDGYVI KVENGYVYVL 140  
 KPGSKRKNIIR TKQQIAEQVA KGTKEAKEKG LAQVAHLSKE EVAAVNEAKR QGRYTTDDGY IFSPTDIIDD 210  
 LGDAYLVPHG NHYHYIPKDD LSPSELAAAQ AYWSQKQGRG ARPSDYRPTP AGRRKAPIP DVTNPGQGH 280  
 QPDNGGYHPA PPRPNDSQN KHQRDEFKKG ELADRYLAGQ TDDNDSGSDH SKPSDKVTH TFLGHRKAY GKGLDGKPYD 350  
 VPHGDHYHII PRSQLSPLEM SIHSDVDSGV TAKHGDHFHY IGFELEQYE LDEVANWKA KGQADELVAA LDQEQGKEKP 420  
 TSDAYVFSKE VTKDGKVGVI MPKDGKOYFY ARYQLDLTQI AFAEQELMLK DKKHYRYDIV DTGIEPRLAV 490  
 LFDTKKVSRL NATYDTGSSF VIPHIDHIHV VPYSWLTRNQ IATIKYVMQH PEVRPDVWSK PGHEESGSI 560  
 DVSSLPMHAG GMPNWIIS AEEVQKALAE GRFAAPDGYI FDPDVLAKE TFVWKDGSFS IPRADGSSLR 630  
 PNVTPLDKRA EWQQAQELLA KKNAGDATDT DKPEEKQAD KSNENQQPSE ASKEEKESDD FIDSLPDYGL 700  
 TINKSDLSQA DRATLEDHIN QLAQKANIDP KYLIFQPEGV QFYNKNGELV TYDIKTLQOI NP 770  
 822

## SEQUENCE LISTING

<110> Heinrichs, Jon  
Johnson, Leslie S.  
Koenig, Scott  
Adamou, John E.

<120> Pneumococcal Protein Homologs and Fragments for  
Vaccines

<130> 469201-402

<140>

<141>

<150> U.S. 60/150,750

<151> 1999-08-25

<160> 6

<170> PatentIn Ver. 2.1

<210> 1

<211> 2478

<212> DNA

<213> Streptococcus pyogenes

<400> 1

```
gtgaagaaaa catatgggta tatcggtca gttgctgcta ttttactagc tactcatatt 60
ggaagttacc aacttggtaa gcatcatatg ggttcagcaa caaaggacaa tcaaattgcc 120
tatattgatg atagcaaagg taaggcaaaa gcccctaaaa caaacaaaac gatggatcaa 180
atcagtgtcg aagaaggcat ctctgtgtaa cagatcgtag tcaaaattac tgaccaaggc 240
tatgtgacct cacatgggtga ccattatcat ttttacaatg ggaaagtcc ttatgatgcg 300
attattagtg aagagtgtgt gatgacggat cctaattacc gttttaaaca atcagacgtt 360
atcaatgaaa tcttagacgg ttacgttatt aaagtcaatg gcaactatta tgtttacctc 420
aagccaggta gcaagcgcaa aaacattcga accaaacaac aaattgctga gcaagtagcc 480
aaaggaacta aagaagctaa agaaaaaggt ttagctcaag tggcccatct cagtaaagaa 540
gaagttgcgg cagtcaatga agcaaaaaga caaggacgct atactacaga cgatggctat 600
atTTTTtagtc cgacagatat cattgatgat ttaggagatg cttatttagt acctcatggt 660
aatcactatc attatatcc taaaaaggat ttgtctccaa gtgagctagc tgctgcacaa 720
gcctactgga gtcaaaaaca aggtcgaggt gctagaccgt ctgattaccg cccgacacca 780
gccccagccc caggtcgtag gaaagcccca attcctgatg tgacgcctaa ccctggacaa 840
ggtcacagc cagataacgg tggctatcat ccagcgcctc ctaggccaaa tgatgcgtca 900
caaaacaaac accaaagaga tgagtttaaa ggaaaaacct ttaaggaact tttagatcaa 960
ctacaccgtc ttgatttgaa ataccgtcat gtggaagaag atgggttgat ttttgaaccg 1020
actcaagtga tcaaatcaaa cgcttttggg tatgtggtgc ctcatggaga tcattatcat 1080
attatcccaa gaagtcagtt atcacctctt gaaatggaat tagcagatcg atacttagcc 1140
ggccaaactg aggacgatga ctcaggttca gatcactcaa aaccatcaga taaagaagtg 1200
acacatacct ttcttgggtca tcgcatcaaa gcttacggaa aaggcttaga tggtaaacca 1260
tatgatacga gtgatgctta tgTTTTtagt aaagaatcca ttcattcagt ggataaatca 1320
ggagttacag ctaaacacgg agatcatttc cactatatag gatttggaga acttgaacaa 1380
tatgagttgg atgaggtcgc taactgggtg aaagcaaaag gtcaagctga tgagcttgct 1440
gctgctttgg atcaggaaca aggcaaagaa aaaccactct ttgacactaa aaaagttagt 1500
cgcaaagtaa caaaagatgg taaagtgggc tatatgatgc caaaagatgg caaggactat 1560
ttctatgtc gtgatcaact tgatttgact cagattgcct ttgccgaaca agaactaatg 1620
cttaaagata agaaacatta ccgttatgac attgttgaca caggtattga gccacgactt 1680
gctgtagatg tgtcaagtct gccgatgcat gctggtaatg ctacttacga tactggaagt 1740
tcgtttgtta tccctcatat tgatcatatc catgtcgttc cgtattcatg gttgacgcgc 1800
gatcagattg caacaatcaa gtatgtgatg caacaccccg aagttcgtcc ggatatatgg 1860
```

```

tctaagccag ggcataaaga gtcagggttcg gtcattccaa atggttacgcc tcttgataaa 1920
cgtgctggta tgccaaactg gcaaattatc cattctgctg aagaagttca aaaagcccta 1980
gcagaaggtc gttttgcaac accagacggc tatattttcg atccacgaga tgttttggcc 2040
aaagaaactt ttgtatggaa agatggctcc tttagcatcc caagagcaga tggcagttca 2100
ttgagaacca ttaataaatc tgatctatcc caagctgagt ggcaacaagc tcaagagtta 2160
ttggcaaaga aaaacgctgg tgatgctact gatacggata aacccaaaga aaagcaacag 2220
gcagataaga gcaatgaaaa ccaacagcca agtgaagcca gtaaagaaga agaaaaagaa 2280
tcagatgact ttatagacag tttaccagac tatgggtctag atagagcaac cctagaagat 2340
catatcaatc aattagcaca aaaagcta atcgatccta agtatctcat tttccaacca 2400
gaagggtgcc aattttataa taaaaatggg gaattggtaa cttatgatat caagacactt 2460
caacaaataa acccttaa

```

&lt;210&gt; 2

&lt;211&gt; 825

&lt;212&gt; PRT

&lt;213&gt; Streptococcus pyogenes

&lt;400&gt; 2

```

Val Lys Lys Thr Tyr Gly Tyr Ile Gly Ser Val Ala Ala Ile Leu Leu
  1             5             10             15

Ala Thr His Ile Gly Ser Tyr Gln Leu Gly Lys His His Met Gly Ser
      20             25             30

Ala Thr Lys Asp Asn Gln Ile Ala Tyr Ile Asp Asp Ser Lys Gly Lys
      35             40             45

Ala Lys Ala Pro Lys Thr Asn Lys Thr Met Asp Gln Ile Ser Ala Glu
      50             55             60

Glu Gly Ile Ser Ala Glu Gln Ile Val Val Lys Ile Thr Asp Gln Gly
      65             70             75             80

Tyr Val Thr Ser His Gly Asp His Tyr His Phe Tyr Asn Gly Lys Val
      85             90             95

Pro Tyr Asp Ala Ile Ile Ser Glu Glu Leu Leu Met Thr Asp Pro Asn
      100            105            110

Tyr Arg Phe Lys Gln Ser Asp Val Ile Asn Glu Ile Leu Asp Gly Tyr
      115            120            125

Val Ile Lys Val Asn Gly Asn Tyr Tyr Val Tyr Leu Lys Pro Gly Ser
      130            135            140

Lys Arg Lys Asn Ile Arg Thr Lys Gln Gln Ile Ala Glu Gln Val Ala
      145            150            155            160

Lys Gly Thr Lys Glu Ala Lys Glu Lys Gly Leu Ala Gln Val Ala His
      165            170            175

Leu Ser Lys Glu Glu Val Ala Ala Val Asn Glu Ala Lys Arg Gln Gly
      180            185            190

Arg Tyr Thr Thr Asp Asp Gly Tyr Ile Phe Ser Pro Thr Asp Ile Ile
      195            200            205

Asp Asp Leu Gly Asp Ala Tyr Leu Val Pro His Gly Asn His Tyr His

```

210		215		220															
Tyr 225	Ile	Pro	Lys	Lys	Asp 230	Leu	Ser	Pro	Ser	Glu 235	Leu	Ala	Ala	Ala	Gln 240				
Ala	Tyr	Trp	Ser	Gln 245	Lys	Gln	Gly	Arg	Gly 250	Ala	Arg	Pro	Ser	Asp 255	Tyr				
Arg	Pro	Thr	Pro 260	Ala	Pro	Ala	Pro	Gly 265	Arg	Arg	Lys	Ala	Pro 270	Ile	Pro				
Asp	Val 275	Thr	Pro	Asn	Pro	Gly	Gln 280	Gly	His	Gln	Pro	Asp 285	Asn	Gly	Gly				
Tyr 290	His	Pro	Ala	Pro	Pro	Arg 295	Pro	Asn	Asp	Ala	Ser 300	Gln	Asn	Lys	His				
Gln 305	Arg	Asp	Glu	Phe	Lys 310	Gly	Lys	Thr	Phe	Lys 315	Glu	Leu	Leu	Asp	Gln				
Leu	His	Arg	Leu	Asp 325	Leu	Lys	Tyr	Arg	His 330	Val	Glu	Glu	Asp	Gly 335	Leu				
Ile	Phe	Glu	Pro 340	Thr	Gln	Val	Ile	Lys 345	Ser	Asn	Ala	Phe	Gly 350	Tyr	Val				
Val	Pro	His 355	Gly	Asp	His	Tyr	His 360	Ile	Ile	Pro	Arg 365	Ser	Gln	Leu	Ser				
Pro	Leu	Glu	Met	Glu	Leu	Ala 375	Asp	Arg	Tyr	Leu	Ala 380	Gly	Gln	Thr	Glu				
Asp 385	Asp	Asp	Ser	Gly	Ser 390	Asp	His	Ser	Lys	Pro 395	Ser	Asp	Lys	Glu	Val 400				
Thr	His	Thr	Phe	Leu 405	Gly	His	Arg	Ile	Lys 410	Ala	Tyr	Gly	Lys	Gly 415	Leu				
Asp	Gly	Lys	Pro	Tyr	Asp	Thr	Ser	Asp 425	Ala	Tyr	Val	Phe	Ser 430	Lys	Glu				
Ser	Ile	His 435	Ser	Val	Asp	Lys	Ser 440	Gly	Val	Thr	Ala	Lys 445	His	Gly	Asp				
His 450	Phe	His	Tyr	Ile	Gly	Phe 455	Gly	Glu	Leu	Glu	Gln 460	Tyr	Glu	Leu	Asp				
Glu 465	Val	Ala	Asn	Trp	Val	Lys	Ala	Lys	Gly	Gln 475	Ala	Asp	Glu	Leu	Ala 480				
Ala	Ala	Leu	Asp	Gln 485	Glu	Gln	Gly	Lys	Glu 490	Lys	Pro	Leu	Phe	Asp 495	Thr				
Lys	Lys	Val	Ser	Arg	Lys	Val	Thr	Lys 505	Asp	Gly	Lys	Val 510	Gly	Tyr	Met				
Met	Pro	Lys	Asp	Gly	Lys	Asp	Tyr	Phe 520	Tyr	Ala	Arg	Asp 525	Gln	Leu	Asp				

Leu Thr Gln Ile Ala Phe Ala Glu Gln Glu Leu Met Leu Lys Asp Lys  
 530 535 540  
 Lys His Tyr Arg Tyr Asp Ile Val Asp Thr Gly Ile Glu Pro Arg Leu  
 545 550 555 560  
 Ala Val Asp Val Ser Ser Leu Pro Met His Ala Gly Asn Ala Thr Tyr  
 565 570 575  
 Asp Thr Gly Ser Ser Phe Val Ile Pro His Ile Asp His Ile His Val  
 580 585 590  
 Val Pro Tyr Ser Trp Leu Thr Arg Asp Gln Ile Ala Thr Ile Lys Tyr  
 595 600 605  
 Val Met Gln His Pro Glu Val Arg Pro Asp Ile Trp Ser Lys Pro Gly  
 610 615 620  
 His Glu Glu Ser Gly Ser Val Ile Pro Asn Val Thr Pro Leu Asp Lys  
 625 630 635 640  
 Arg Ala Gly Met Pro Asn Trp Gln Ile Ile His Ser Ala Glu Glu Val  
 645 650 655  
 Gln Lys Ala Leu Ala Glu Gly Arg Phe Ala Thr Pro Asp Gly Tyr Ile  
 660 665 670  
 Phe Asp Pro Arg Asp Val Leu Ala Lys Glu Thr Phe Val Trp Lys Asp  
 675 680 685  
 Gly Ser Phe Ser Ile Pro Arg Ala Asp Gly Ser Ser Leu Arg Thr Ile  
 690 695 700  
 Asn Lys Ser Asp Leu Ser Gln Ala Glu Trp Gln Gln Ala Gln Glu Leu  
 705 710 715 720  
 Leu Ala Lys Lys Asn Ala Gly Asp Ala Thr Asp Thr Asp Lys Pro Lys  
 725 730 735  
 Glu Lys Gln Gln Ala Asp Lys Ser Asn Glu Asn Gln Gln Pro Ser Glu  
 740 745 750  
 Ala Ser Lys Glu Glu Glu Lys Glu Ser Asp Asp Phe Ile Asp Ser Leu  
 755 760 765  
 Pro Asp Tyr Gly Leu Asp Arg Ala Thr Leu Glu Asp His Ile Asn Gln  
 770 775 780  
 Leu Ala Gln Lys Ala Asn Ile Asp Pro Lys Tyr Leu Ile Phe Gln Pro  
 785 790 795 800  
 Glu Gly Val Gln Phe Tyr Asn Lys Asn Gly Glu Leu Val Thr Tyr Asp  
 805 810 815  
 Ile Lys Thr Leu Gln Gln Ile Asn Pro  
 820 825

&lt;210&gt; 3

<211> 2379  
 <212> DNA  
 <213> Streptococcus pyogenes

<400> 3  
 atgaaaacga aaaaagttat tatttttagtt ggtctattgt tatcatctca gttgactttg 60  
 atagcttgtc aatcacgagg taatgggtaca tatcccatta aaacgaaaca atcacgtaag 120  
 ggaatgacgt caaacaaaat taaaccgatt aaaaaaagca aaaagacaaa caagactcac 180  
 aaagggtgtgg cgggtgtcga ttttcctaca gatgatgggt ttattttaac caaagactca 240  
 aaaatccttat caaaaacaga tcaggggaatc gttgttgacc atgatgggtca ttcgcatttt 300  
 attttttatg ccgattttaa gggaagtcca tttgaatacc ttattccaaa aggagcaagt 360  
 ttagctaagc cagctggtgc tcagcgagca gctagtcaag ggacttctaa agtagcagat 420  
 cctcatcacc attatgaatt taaccacgag gatattgtgg ctgaagatgc tttaggctac 480  
 acggttcgcc acgatgatca cttccattat attttgaagt caagcttatc aggtcagaca 540  
 caggcacaag ctaaacagggt tgctactcgc ttgccacaaa ccagtagcct tgtttcaaca 600  
 gctacagcta atggtattcc aggcttgcat ttcccaacct cagatgggtt tcaatttaac 660  
 ggtcaaggta ttgttggggg aacaaaagac agtatttttag tggaccacga tggtcactta 720  
 catcctatct cttttgcgga ccttcgtcag ggtggctggg cacatgtggc agatcaatac 780  
 gatcccgcta aaaaagcaga aaagccagca gaaacccatc agacaccaga gctatctgaa 840  
 cgtgaaaagg aataccaaga aaaattagct tattttggcag aaaaattggg gattgatcca 900  
 tcaactatta aacgtgtgga aacacaagac ggtaaacttg gtttggataa ccctcaccat 960  
 gaccacgcac acgtattgat gttatctgat attgaaatcg gaaaagacat tccagatcca 1020  
 catgctattg agcatgcccg tgaattggaa aaacataagg ttggaatgga taccttgccg 1080  
 gccttagggg ttgatgaaga agtgattttg gatatcggtc gcactcacga tgcctcaacc 1140  
 ccattcccat caaatgaaaa agatccgaat atgatgaaag aatgggttagc aacgggttatc 1200  
 aaacttgact tgggcagccg taaagatcct ttgcaacgta aaggactttc actgtttacc 1260  
 aacttagaaa ctttaggaat tggctttaca ccaatcaaag atatctcacc tgttttgcaa 1320  
 tttaaaaaat tgaaacagtt gttaatgaca aaaacagggg tgactgatta tagatttttg 1380  
 gataatatgc cacagttaga aggcattgat atttcacaaa acaatctcaa agatattagt 1440  
 ttcttgagca aatataaaaa cttaactcta gtagcggtcg ctgataatgg tattgaagat 1500  
 attaggccgc ttggtcaatt accaaatctc aaattcctcg tattgagtaa caataagatt 1560  
 tctgatttaa gccactggc atcgttacat caattgcaag aattgcacat tgataataat 1620  
 cagattacag atttaagccc tgtttctcat aaagaatcat tgacgggtgt tgatttatca 1680  
 agaaatgctg atgttgactt agcaacactt caagcaccca aattagaaac gttaatggtc 1740  
 aatgatacca aggtttctca tttggatttc ttgaaaaata atcctaactc atctagccta 1800  
 tctattaacc gtgcgcaatt gcaatctctt gaagggtattg aagcaagtag cgtcattgtc 1860  
 agagtagaag cagaaggtaa ccaaattaaa tcgcttggtc ttaaagacaa gcaagggtca 1920  
 cttactttct tggatgtgac aggcacaccag ttgacttctc tagaagggtg taataatttt 1980  
 acagcacttg acattttaag cgtgtctaaa aaaccaattaa caaatgtcaa cctatctaaa 2040  
 cccaataaga cagttactaa cattgatatt agtcataaca atatctcatt agcagacctt 2100  
 aaattgaacg agcaacatat tccagaagcc attgcgaaaa acttcccagc ggtttacgaa 2160  
 ggttctatgg taggtaatgg aacagctgaa gaaaaagcag ctatgggtac taaggcgaaa 2220  
 gaaagtgtct aagaagcatc ggaatcacat gactacaacc ataatcatal ctatgaagat 2280  
 gaagaagggtc atgctcacga gcacagagac aaagatgatc acgaccatga acatgaggat 2340  
 gaaaatgaag ctaaagatga gcaaaacctat gctgactaa 2379

<210> 4  
 <211> 792  
 <212> PRT  
 <213> Streptococcus pyogenes

<400> 4  
 Met Lys Thr Lys Lys Val Ile Ile Leu Val Gly Leu Leu Leu Ser Ser  
 1 5 10 15  
 Gln Leu Thr Leu Ile Ala Cys Gln Ser Arg Gly Asn Gly Thr Tyr Pro  
 20 25 30

```

Ile Lys Thr Lys Gln Ser Arg Lys Gly Met Thr Ser Asn Lys Ile Lys
  35                                40                                45

Pro Ile Lys Lys Ser Lys Lys Thr Asn Lys Thr His Lys Gly Val Ala
  50                                55                                60

Gly Val Asp Phe Pro Thr Asp Asp Gly Phe Ile Leu Thr Lys Asp Ser
  65                                70                                75                                80

Lys Ile Leu Ser Lys Thr Asp Gln Gly Ile Val Val Asp His Asp Gly
  85                                90                                95

His Ser His Phe Ile Phe Tyr Ala Asp Leu Lys Gly Ser Pro Phe Glu
 100                                105                                110

Tyr Leu Ile Pro Lys Gly Ala Ser Leu Ala Lys Pro Ala Val Ala Gln
 115                                120                                125

Arg Ala Ala Ser Gln Gly Thr Ser Lys Val Ala Asp Pro His His His
 130                                135                                140

Tyr Glu Phe Asn Pro Ala Asp Ile Val Ala Glu Asp Ala Leu Gly Tyr
 145                                150                                155                                160

Thr Val Arg His Asp Asp His Phe His Tyr Ile Leu Lys Ser Ser Leu
 165                                170                                175

Ser Gly Gln Thr Gln Ala Gln Ala Lys Gln Val Ala Thr Arg Leu Pro
 180                                185                                190

Gln Thr Ser Ser Leu Val Ser Thr Ala Thr Ala Asn Gly Ile Pro Gly
 195                                200                                205

Leu His Phe Pro Thr Ser Asp Gly Phe Gln Phe Asn Gly Gln Gly Ile
 210                                215                                220

Val Gly Val Thr Lys Asp Ser Ile Leu Val Asp His Asp Gly His Leu
 225                                230                                235                                240

His Pro Ile Ser Phe Ala Asp Leu Arg Gln Gly Gly Trp Ala His Val
 245                                250                                255

Ala Asp Gln Tyr Asp Pro Ala Lys Lys Ala Glu Lys Pro Ala Glu Thr
 260                                265                                270

His Gln Thr Pro Glu Leu Ser Glu Arg Glu Lys Glu Tyr Gln Glu Lys
 275                                280                                285

Leu Ala Tyr Leu Ala Glu Lys Leu Gly Ile Asp Pro Ser Thr Ile Lys
 290                                295                                300

Arg Val Glu Thr Gln Asp Gly Lys Leu Gly Leu Glu Tyr Pro His His
 305                                310                                315                                320

Asp His Ala His Val Leu Met Leu Ser Asp Ile Glu Ile Gly Lys Asp
 325                                330                                335

Ile Pro Asp Pro His Ala Ile Glu His Ala Arg Glu Leu Glu Lys His
 340                                345                                350

```

Lys Val Gly Met Asp Thr Leu Arg Ala Leu Gly Phe Asp Glu Glu Val  
 355 360 365  
 Ile Leu Asp Ile Val Arg Thr His Asp Ala Pro Thr Pro Phe Pro Ser  
 370 375 380  
 Asn Glu Lys Asp Pro Asn Met Met Lys Glu Trp Leu Ala Thr Val Ile  
 385 390 395 400  
 Lys Leu Asp Leu Gly Ser Arg Lys Asp Pro Leu Gln Arg Lys Gly Leu  
 405 410 415  
 Ser Leu Leu Pro Asn Leu Glu Thr Leu Gly Ile Gly Phe Thr Pro Ile  
 420 425 430  
 Lys Asp Ile Ser Pro Val Leu Gln Phe Lys Lys Leu Lys Gln Leu Leu  
 435 440 445  
 Met Thr Lys Thr Gly Val Thr Asp Tyr Arg Phe Leu Asp Asn Met Pro  
 450 455 460  
 Gln Leu Glu Gly Ile Asp Ile Ser Gln Asn Asn Leu Lys Asp Ile Ser  
 465 470 475 480  
 Phe Leu Ser Lys Tyr Lys Asn Leu Thr Leu Val Ala Ala Ala Asp Asn  
 485 490 495  
 Gly Ile Glu Asp Ile Arg Pro Leu Gly Gln Leu Pro Asn Leu Lys Phe  
 500 505 510  
 Leu Val Leu Ser Asn Asn Lys Ile Ser Asp Leu Ser Pro Leu Ala Ser  
 515 520 525  
 Leu His Gln Leu Gln Glu Leu His Ile Asp Asn Asn Gln Ile Thr Asp  
 530 535 540  
 Leu Ser Pro Val Ser His Lys Glu Ser Leu Thr Val Val Asp Leu Ser  
 545 550 555 560  
 Arg Asn Ala Asp Val Asp Leu Ala Thr Leu Gln Ala Pro Lys Leu Glu  
 565 570 575  
 Thr Leu Met Val Asn Asp Thr Lys Val Ser His Leu Asp Phe Leu Lys  
 580 585 590  
 Asn Asn Pro Asn Leu Ser Ser Leu Ser Ile Asn Arg Ala Gln Leu Gln  
 595 600 605  
 Ser Leu Glu Gly Ile Glu Ala Ser Ser Val Ile Val Arg Val Glu Ala  
 610 615 620  
 Glu Gly Asn Gln Ile Lys Ser Leu Val Leu Lys Asp Lys Gln Gly Ser  
 625 630 635 640  
 Leu Thr Phe Leu Asp Val Thr Gly Asn Gln Leu Thr Ser Leu Glu Gly  
 645 650 655  
 Val Asn Asn Phe Thr Ala Leu Asp Ile Leu Ser Val Ser Lys Asn Gln

660                      665                      670  
 Leu Thr Asn Val Asn Leu Ser Lys Pro Asn Lys Thr Val Thr Asn Ile  
           675                      680                      685  
 Asp Ile Ser His Asn Asn Ile Ser Leu Ala Asp Leu Lys Leu Asn Glu  
           690                      695                      700  
 Gln His Ile Pro Glu Ala Ile Ala Lys Asn Phe Pro Ala Val Tyr Glu  
           705                      710                      715                      720  
 Gly Ser Met Val Gly Asn Gly Thr Ala Glu Glu Lys Ala Ala Met Ala  
                                  725                      730                      735  
 Thr Lys Ala Lys Glu Ser Ala Gln Glu Ala Ser Glu Ser His Asp Tyr  
                                  740                      745                      750  
 Asn His Asn His Thr Tyr Glu Asp Glu Glu Gly His Ala His Glu His  
                                  755                      760                      765  
 Arg Asp Lys Asp Asp His Asp His Glu His Glu Asp Glu Asn Glu Ala  
                                  770                      775                      780  
 Lys Asp Glu Gln Asn His Ala Asp  
                                  785                      790

&lt;210&gt; 5

&lt;211&gt; 2469

&lt;212&gt; DNA

&lt;213&gt; Streptococcus agalactiae

&lt;400&gt; 5

gtgaagaaaa catatgggta tatcggtcga gttgctgcta ttttactagc tactcatatt 60  
 ggaagttacc agcttggtta gcatcatatg ggtctagcaa caaaggacaa tcagattgcc 120  
 tatattgatg atagcaaagg taaggtaaaa gcccttaaaa caaacaaaac gatggatcaa 180  
 atcagtgctg aagaaggcat ctctgctgaa cagatcgtag tcaaaattac tgaccaaggt 240  
 tatgtttacct cacacgggtga ccattatcat ttttacaatg ggaaagttcc ttatgatgcg 300  
 attattagtg aagagttggt gatgacggat cctaattacc atttttaaca atcagacggt 360  
 atcaatgaaa tcttagacgg ttacgttatt aaagtcaatg gcaactatta tgtttacctc 420  
 aagccaggta gtaagcgcaa aaacattcga accaaacaac aaattgctga gcaagtagcc 480  
 aaaggaacta aagaagctaa agaaaaaggt ttagctcaag tggcccatct cagtaaagaa 540  
 gaagttgcgg cagtcaatga agcaaaaaga caaggacgct atactacaga cgatggctat 600  
 atttttagtc cgacagatat cattgatgat ttaggagatg cttatttagt acctcatggt 660  
 aatcactatc atttatattcc taaaaaagat ttgtctccaa gtgagctagc tgctgcacaa 720  
 gcctactgga gtcaaaaaca aggtcgaggt gctagaccgt ctgattaccg cccgacacca 780  
 gccccaggtc gtaggaaagc cccaattcct gatgtgacgc ctaaccctgg acaaggtcat 840  
 cagccagata acggtgggta tcatccagcg cctcctaggc caaatgatgc gtcacaaaac 900  
 aaacacccaaa gagatgagtt taaaggaaaa acctttaagg aactttttaga tcaactacac 960  
 cgtcttgatt tgaaataaccg tcatgtggaa gaagatgggt tgatttttga accgactcaa 1020  
 gtgatcaaat caaacgcttt tgggtatgtg gtgcctcatg gagatcatta tcatattatc 1080  
 ccaagaagtc agttatcacc acttgaaatg gaattagcag atcgatactt agccggccaa 1140  
 actgatgaca acgactcagg ttcagatcac tcaaaacat cagataaaga agtgacacat 1200  
 acctttcttg gtcacgcat caaagcttac ggaaaaggct tagatggtaa accatatgat 1260  
 acgagtgatg cttatgtttt tagtaaagaa tccattcatt cagtggataa atcaggagtt 1320  
 acagctaaac acggagatca tttccactat ataggatttg gagaacttga acaatatgag 1380  
 ttggtgagg tcgctaactg ggtgaaagca aaagggtcaag ctgatgagct tgttgctgct 1440  
 ttggatcagg aacaaggcaa agaaaaacca ctctttgaca ctaaaaaagt gagtgcgcaa 1500  
 gtaacaaaaag atggtaaagt gggctatatt atgccaaaag atggcaagga ctatttctat 1560

```

gctcgttatt aacttgattt gactcagatt gcctttgccg aacaagaact aatgcttaaa 1620
gataagaagc attaccgtta tgacattggt gatacaggca ttgagccacg acttgctgta 1680
gatgtgtcaa gtctgccgat gcatgctggt aatgctactt acgatactgg aagttcgttt 1740
gttatccccc atattgatca tatccatgtc gttccgtatt catggttgac gcgcaatcag 1800
attgcaacaa tcaagtatgt gatgcaacac cccgaagtgc gtccggatgt atgggtctaag 1860
ccagggcatg aagagtcagg ttcggtcatt ccaaagtta cgctcttga taaacgtgct 1920
ggtatgccaa actggcaaat tatccattct gctgaagaag ttcaaaaagc cctagcagaa 1980
ggtcgttttg cagcaccaga cggctatatt ttcgatccac gagatgtttt ggcaaaaagaa 2040
acttttgat ggaaagatgg ctcccttagc atcccaagag cagatggcag ttcattgaga 2100
accattaata aatccgatct atcccaagct gagtggcaac aagctcaaga gttattggca 2160
aagaaaaatg ctggtgatgc tactgatacg gataaacctg aagaaaaagca acaggcagat 2220
aagagcaatg aaaaccaaca gccaagtga gccagtaaag aagaaaaaga atcagatgac 2280
tttatagaca gtttaccaga ctatggtcta gatagagcaa ccctagaaga tcatatcaat 2340
caattagcac aaaaagctaa tatcgatcct aagtatctca ttttccaacc agaaggtgtc 2400
caattttata ataaaaatgg tgaattggta acttatgata tcaagacact tcaacaaata 2460
aaccccttaa 2469

```

&lt;210&gt; 6

&lt;211&gt; 822

&lt;212&gt; PRT

&lt;213&gt; Streptococcus agalactiae

&lt;400&gt; 6

```

Val Lys Lys Thr Tyr Gly Tyr Ile Gly Ser Val Ala Ala Ile Leu Leu
  1             5             10             15

```

```

Ala Thr His Ile Gly Ser Tyr Gln Leu Gly Lys His His Met Gly Leu
          20             25             30

```

```

Ala Thr Lys Asp Asn Gln Ile Ala Tyr Ile Asp Asp Ser Lys Gly Lys
          35             40             45

```

```

Val Lys Ala Pro Lys Thr Asn Lys Thr Met Asp Gln Ile Ser Ala Glu
          50             55             60

```

```

Glu Gly Ile Ser Ala Glu Gln Ile Val Val Lys Ile Thr Asp Gln Gly
          65             70             75             80

```

```

Tyr Val Thr Ser His Gly Asp His Tyr His Phe Tyr Asn Gly Lys Val
          85             90             95

```

```

Pro Tyr Asp Ala Ile Ile Ser Glu Glu Leu Leu Met Thr Asp Pro Asn
          100            105            110

```

```

Tyr His Phe Lys Gln Ser Asp Val Ile Asn Glu Ile Leu Asp Gly Tyr
          115            120            125

```

```

Val Ile Lys Val Asn Gly Asn Tyr Tyr Val Tyr Leu Lys Pro Gly Ser
          130            135            140

```

```

Lys Arg Lys Asn Ile Arg Thr Lys Gln Gln Ile Ala Glu Gln Val Ala
          145            150            155            160

```

```

Lys Gly Thr Lys Glu Ala Lys Glu Lys Gly Leu Ala Gln Val Ala His
          165            170            175

```

```

Leu Ser Lys Glu Glu Val Ala Ala Val Asn Glu Ala Lys Arg Gln Gly
          180            185            190

```

Arg Tyr Thr Thr Asp Asp Gly Tyr Ile Phe Ser Pro Thr Asp Ile Ile  
 195 200 205  
 Asp Asp Leu Gly Asp Ala Tyr Leu Val Pro His Gly Asn His Tyr His  
 210 215 220  
 Tyr Ile Pro Lys Lys Asp Leu Ser Pro Ser Glu Leu Ala Ala Ala Gln  
 225 230 235 240  
 Ala Tyr Trp Ser Gln Lys Gln Gly Arg Gly Ala Arg Pro Ser Asp Tyr  
 245 250 255  
 Arg Pro Thr Pro Ala Pro Gly Arg Arg Lys Ala Pro Ile Pro Asp Val  
 260 265 270  
 Thr Pro Asn Pro Gly Gln Gly His Gln Pro Asp Asn Gly Gly Tyr His  
 275 280 285  
 Pro Ala Pro Pro Arg Pro Asn Asp Ala Ser Gln Asn Lys His Gln Arg  
 290 295 300  
 Asp Glu Phe Lys Gly Lys Thr Phe Lys Glu Leu Leu Asp Gln Leu His  
 305 310 315 320  
 Arg Leu Asp Leu Lys Tyr Arg His Val Glu Glu Asp Gly Leu Ile Phe  
 325 330 335  
 Glu Pro Thr Gln Val Ile Lys Ser Asn Ala Phe Gly Tyr Val Val Pro  
 340 345 350  
 His Gly Asp His Tyr His Ile Ile Pro Arg Ser Gln Leu Ser Pro Leu  
 355 360 365  
 Glu Met Glu Leu Ala Asp Arg Tyr Leu Ala Gly Gln Thr Asp Asp Asn  
 370 375 380  
 Asp Ser Gly Ser Asp His Ser Lys Pro Ser Asp Lys Glu Val Thr His  
 385 390 395 400  
 Thr Phe Leu Gly His Arg Ile Lys Ala Tyr Gly Lys Gly Leu Asp Gly  
 405 410 415  
 Lys Pro Tyr Asp Thr Ser Asp Ala Tyr Val Phe Ser Lys Glu Ser Ile  
 420 425 430  
 His Ser Val Asp Lys Ser Gly Val Thr Ala Lys His Gly Asp His Phe  
 435 440 445  
 His Tyr Ile Gly Phe Gly Glu Leu Glu Gln Tyr Glu Leu Asp Glu Val  
 450 455 460  
 Ala Asn Trp Val Lys Ala Lys Gly Gln Ala Asp Glu Leu Val Ala Ala  
 465 470 475 480  
 Leu Asp Gln Glu Gln Gly Lys Glu Lys Pro Leu Phe Asp Thr Lys Lys  
 485 490 495  
 Val Ser Arg Lys Val Thr Lys Asp Gly Lys Val Gly Tyr Ile Met Pro

500					505					510						
Lys	Asp	Gly	Lys	Asp	Tyr	Phe	Tyr	Ala	Arg	Tyr	Gln	Leu	Asp	Leu	Thr	
515					520					525						
Gln	Ile	Ala	Phe	Ala	Glu	Gln	Glu	Leu	Met	Leu	Lys	Asp	Lys	Lys	His	
530					535					540						
Tyr	Arg	Tyr	Asp	Ile	Val	Asp	Thr	Gly	Ile	Glu	Pro	Arg	Leu	Ala	Val	
545					550					555					560	
Asp	Val	Ser	Ser	Leu	Pro	Met	His	Ala	Gly	Asn	Ala	Thr	Tyr	Asp	Thr	
565					570					575						
Gly	Ser	Ser	Phe	Val	Ile	Pro	His	Ile	Asp	His	Ile	His	Val	Val	Pro	
580					585					590						
Tyr	Ser	Trp	Leu	Thr	Arg	Asn	Gln	Ile	Ala	Thr	Ile	Lys	Tyr	Val	Met	
595					600					605						
Gln	His	Pro	Glu	Val	Arg	Pro	Asp	Val	Trp	Ser	Lys	Pro	Gly	His	Glu	
610					615					620						
Glu	Ser	Gly	Ser	Val	Ile	Pro	Asn	Val	Thr	Pro	Leu	Asp	Lys	Arg	Ala	
625					630					635					640	
Gly	Met	Pro	Asn	Trp	Gln	Ile	Ile	His	Ser	Ala	Glu	Glu	Val	Gln	Lys	
645					650					655						
Ala	Leu	Ala	Glu	Gly	Arg	Phe	Ala	Ala	Pro	Asp	Gly	Tyr	Ile	Phe	Asp	
660					665					670						
Pro	Arg	Asp	Val	Leu	Ala	Lys	Glu	Thr	Phe	Val	Trp	Lys	Asp	Gly	Ser	
675					680					685						
Phe	Ser	Ile	Pro	Arg	Ala	Asp	Gly	Ser	Ser	Leu	Arg	Thr	Ile	Asn	Lys	
690					695					700						
Ser	Asp	Leu	Ser	Gln	Ala	Glu	Trp	Gln	Gln	Ala	Gln	Glu	Leu	Leu	Ala	
705					710					715					720	
Lys	Lys	Asn	Ala	Gly	Asp	Ala	Thr	Asp	Thr	Asp	Lys	Pro	Glu	Glu	Lys	
725					730					735						
Gln	Gln	Ala	Asp	Lys	Ser	Asn	Glu	Asn	Gln	Gln	Pro	Ser	Glu	Ala	Ser	
740					745					750						
Lys	Glu	Glu	Lys	Glu	Ser	Asp	Asp	Phe	Ile	Asp	Ser	Leu	Pro	Asp	Tyr	
755					760					765						
Gly	Leu	Asp	Arg	Ala	Thr	Leu	Glu	Asp	His	Ile	Asn	Gln	Leu	Ala	Gln	
770					775					780						
Lys	Ala	Asn	Ile	Asp	Pro	Lys	Tyr	Leu	Ile	Phe	Gln	Pro	Glu	Gly	Val	
785					790					795					800	
Gln	Phe	Tyr	Asn	Lys	Asn	Gly	Glu	Leu	Val	Thr	Tyr	Asp	Ile	Lys	Thr	
805					810					815						

WO 01/14421

PCT/US00/23417

Leu Gln Gln Ile Asn Pro  
820

## INTERNATIONAL SEARCH REPORT

Intern. nat. Application No

PCT/US 00/23417

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C07K14/315 C12N15/31 C07K16/12 A61K39/09 A61K39/40

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C07K C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, BIOSIS

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 98 18930 A (HUMAN GENOME SCIENCES INC ;CHOI GIL H (US); HROMOCKYJ ALEX (US); J) 7 May 1998 (1998-05-07) SEQ ID NO:55 page 59	8,9, 11-24
X	SPELLERBERG B ET AL: "Streptococcus agalactiae Lmb (lmb) gene, complete cds;and unknown gene" EMBL NUCLEOTIDE SEQU, 11 February 1999 (1999-02-11), XP002125180 cited in the application the whole document	1-5,7-10

☒ Further documents are listed in the continuation of box C.☒ Patent family members are listed in annex.

## \* Special categories of cited documents:

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

\*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

\*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

\*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

\*G\* document member of the same patent family

Date of the actual completion of the international search

9 January 2001

Date of mailing of the international search report

18.01.01

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Bilang, J

## INTERNATIONAL SEARCH REPORT

Intern    nal Application No

PCT/US 00/23417

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	WO 99 42588 A (BIOCHEM VACCINS INC ; BRODEUR BERNARD R (CA); CHARLEBOIS ISABELLE () 26 August 1999 (1999-08-26) figures 3A, 3C ---	1-3, 5, 8, 9, 11-24
P,X	WO 00 06736 A (HANNIFFY SEAN BOSCO ; LE PAGE RICHARD WILLIAM FALLA (GB); WELLS JER) 10 February 2000 (2000-02-10) clone 18 figure 1 -----	1-3, 5, 7-9, 11-24

# INTERNATIONAL SEARCH REPORT

International application No.  
PCT/US 00/23417

## Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☒ Claims Nos.:  
because they relate to subject matter not required to be searched by this Authority, namely:  
  
Although claims 20-24 are directed to a method of treatment of the human/animal body and/or a diagnostic method practised on the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.
2. ☐ Claims Nos.:  
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
3. ☐ Claims Nos.:  
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

## Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☒ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest.
- ☐ No protest accompanied the payment of additional search fees.

**FURTHER INFORMATION CONTINUED FROM PCT/SA/ 210**

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-24, all partially

Polypeptide having SEQ ID NO: 2 and gene encoding the polynucleotide (SEQ ID NO: 1); antibodies specific for said polypeptide; vector and cells comprising said polynucleotide; composition comprising said polypeptide and methods making use of said polypeptide or antibodies.

2. Claims: 1-24, all partially

Polypeptide having SEQ ID NO: 4 and gene encoding the polynucleotide (SEQ ID NO: 3); antibodies specific for said polypeptide; vector and cells comprising said polynucleotide; composition comprising said polypeptide and methods making use of said polypeptide or antibodies.

3. Claims: 1-24, all partially

Polypeptide having SEQ ID NO: 6 and gene encoding the polynucleotide (SEQ ID NO: 5); antibodies specific for said polypeptide; vector and cells comprising said polynucleotide; composition comprising said polypeptide and methods making use of said polypeptide or antibodies.

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 00/23417

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9818930 A	07-05-1998	AU 5194598 A	22-05-1998
		AU 6909098 A	22-05-1998
		EP 0942983 A	22-09-1999
		EP 0941335 A	15-09-1999
		WO 9818931 A	07-05-1998
		US 6159469 A	12-12-2000
WO 9942588 A	26-08-1999	AU 2505999 A	06-09-1999
		EP 1054971 A	29-11-2000
		NO 20004161 A	19-10-2000
WO 0006736 A	10-02-2000	NONE	